Journal of
Experimental Psychology

VOL. 27, NO. 1 JULY, 1940

# BRIGHTNESS-CONSTANCY IN UNRECOGNIZED SHADOWS

BY R. B. MACLEOD

*Swarthmore College*

## INTRODUCTION

In the standard experiments on the constancy of color it has been found repeatedly [1] that the phenomena of color constancy are closely related to the perception of illumination. It is as though the perceptual field were split into two essentially different components, one stable component—the object colors—and one variable component—the illumination, which functions as the bearer of the changes introduced into the field. Any procedure, *e.g.*, reduction, which interferes with such a severance of object-color and illumination tends at the same time to destroy the phenomena of color-constancy. Must we say, then, that there is a causal link between the perception of illumination as such and the phenomena of color-constancy? If we consider only examples drawn from situations in everyday life, or the experiments patterned after such observations, we are almost forced to give an affirmative answer. We systematically disregard the chromatic tint of an electric light except, for instance, under such artificial conditions as those of stage lighting which render impossible its perception as illumination. If we consider, however, some of the more 'artificial' examples of color-constancy, we find cases in which exactly analogous phenomena are observable even when there is no perception of a changing or anomalous illu-

[1] *Cf.* Katz (9), Koffka (12), and accompanying references.

mination. In the light of such cases, one's conclusion must be, not that the perception of illumination as such is an essential condition of color-constancy phenomena, but that, to some extent at least, both sets of phenomena are products of certain more fundamental conditions within the visual field. Such a conclusion has already been advanced tentatively by Kaila (**6**), Katona (**7**), MacLeod (**13**), Kardos (**8**), Henneman (**4**), Koffka (**11**, **12**) and Helson (**3**). The experiments to be reported here demonstrate that color-constancy phenomena can be produced which are independent of the perception of illumination, and point the way toward the explanation of some of these phenomena at least in terms of the laws of field organization.

## The Encircled Shadow Experiment

According to Hering's well-known 'spot-shadow' experiment (**5**), the characteristic appearance of a shadow will be destroyed if one draws about it a black line which coincides with the penumbra of the shadow. This encircled shadow becomes darker and assumes the appearance of a surface-color. Using such a shadow as background and a smaller disc as infield, Kardos (**8**) demonstrated that with the encirclement of the shadow the disc appeared, not brighter as one might predict from a superficial consideration of the laws of brightness contrast, but considerably darker. The effect was relatively independent of the size of the field, and could be produced as easily with a white as with a black line. An explanation in terms of contrast exercised directly by the line is thus disposed of. The customary explanation of the effect is that where the shadow is perceived as shadow it functions as a special local illumination, with the result that the disc undergoes a brightening in accordance with the laws of brightness-constancy; whereas after the destruction of the penumbra the shadow is transformed into a surface-color, and the disc is seen as darker but in the brighter illumination of the room.

In the Kardos experiment there can be no questioning the fact that the unencircled shadow appears as a shadow, *i.e.*, as an area of reduced illumination, and that with the addition of

the encircling line the character of shadow is lost and is replaced by an appearance approximating that of surface-color. In order to obtain a measure of the change in brightness induced by the addition of the line, the author repeated the Kardos experiment in the following way: On an upright black cardboard a circular white line, 5 mm thick, was painted, and upon this a circular shadow with a perceptible penumbra of approximately 2.5 cm was focussed in such a way that the penumbra coincided approximately with the line. In spite of the fact that the penumbra necessarily overlapped the line to a certain extent, the Hering effect was perfectly clear. Instead of a shadow, the observers saw without exception an intense, almost liminous black, surrounded by a white line. It was usually described as a 'velvety' or 'sooty' black, and the observers, when questioned, usually assumed that it had been painted, that a piece of velvet had been pasted on the cardboard or that a hole had been cut, through which a velvet surface could be seen. When the cardboard with the line was suddenly removed, and replaced by an equivalent cardboard without any line, the appearance of the shadow changed. It was now seen simply as a circular shadow, considerably brighter than the velvet black of the previous observation, through which the color of the cardboard was clearly apparent.

For purposes of measurement, a medium grey disc, 10 cm in diameter, was now set up 6 cm in front of, and at the center of, the shadow, so that the shadow would form its effective background, and, by means of a neutral color-wheel, equations were made for each of the two appearances of the shadow.[2] Table I gives the equations determined for 10 observers. The values are expressed in terms of degrees of white on the color-wheel, black being the only other component. Column 2 gives the equations for the disc seen against the encircled shadow, Column 3 the equations for the disc seen against the unencircled shadow, and Column 4 the differences in brightness between the two steps. It will be seen that, in spite of the large individual differences customary in color-constancy

[2] The experimental conditions and methods are almost identical with those described in detail in MacLeod (13).

TABLE I

BRIGHTNESS EQUATION FOR MEDIUM-GREY DISC SEEN AGAINST BACKGROUNDS OF ENCIRCLED AND UNENCIRCLED SHADOW; VALUES EXPRESSED IN DEGREES OF WHITE ON THE COLOR-WHEEL

| (1) Obs. | (2) Encircled Shadow Background | (3) Unencircled Shadow Background | (4) Difference |
|---|---|---|---|
| 1 | 135 | 180 | 45 |
| 2 | 108 | 162 | 54 |
| 3 | 63 | 86 | 23 |
| 4 | 81 | 162 | 81 |
| 5 | 63 | 96 | 33 |
| 6 | 77 | 171 | 94 |
| 7 | 117 | 144 | 27 |
| 8 | 63 | 81 | 18 |
| 9 | 72 | 90 | 18 |
| Sum | 779 | 1172 | 393 |
| Mean | 86.6 | 130.2 | 43.7 |
| S.D. | 25.2 | 19.7 | |

experiments, there is in every case a marked increase in brightness.

Neither from a phenomenological nor from a quantitative point of view do these results differ significantly from those of other experiments in the field of color-constancy. Kardos has satisfactorily demonstrated that when we destroy shadow-character by introducing figural contours, we are in principle doing nothing other than replacing the shadow by a physically equivalent surface-color. This, as has been reported many times (**9**), yields the customary differences in brightness. Consequently, what we are comparing in the Kardos experiment is simply a shadow which appears as a shadow with a shadow which appears as a thing. We are still faced with the basic question: Must we have a distinction between thing-appearance and shadow-appearance, *i.e.*, between object-color and impression of illumination, if we are to produce these changes in brightness which we have termed color-constancy phenomena?

The problem obviously demands a further 'artificializing' of the experiment. In a sense, we have already departed far from the conditions of everyday perception. The shadows of everyday life are usually tri-dimensional rather than bi-dimen-

sional; they are usually cast by clearly perceptible objects, standing before light sources which are not concealed; and they are frequently in motion. Under such conditions, as Katz has shown (**9**), color-constancy may be perfect. The fact that, even under the unusual conditions of the laboratory, analogous results may be obtained is a tribute to the basic character of the visual processes involved. The question now is whether the phenomena will continue to appear even under conditions so artificial as to preclude the perception of a special illumination as such.

## Experiment with Unrecognized Shadows

This result was obtained by means of a further modification of the Hering 'spot-shadow' experiment, whereby two shadows, one encircled and one unencircled, were presented side by side in such a way that light-source and shadow-casters were concealed from the observer and the judgment that either one was a shadow was rendered exceedingly difficult.[3]

Figure 1 illustrates the apparatus used in the experiment. Against a uniform black wall (*A*) in a dark room two black cardboard screens ($B_1$, $B_2$—72 × 56 cm) were set up side by side. Behind each screen was a motor with a shaft long enough to project through a hole and permit the rotation of a color-wheel ($C_1$, $C_2$) approximately 1 cm in front of the screen. The spatial separation of color-wheel from screen was too small to be perceptible. Each color-wheel was fitted with variable black and white discs, 10 cm in diameter, and their centers were separated from each other by a distance of 62 cm. The only source of illumination was a 100 watt daylight lamp (*D*), attached to a stand 142 cm in front of the screens and so placed that the two screens received exactly equal amounts of light. Between the lamp and the screens were placed two circular metal discs ($E_1$, $E_2$), each attached by a thin wire to a stand in such a way that it cast a round shadow on the corresponding screen. The shadow was in each case centered about the rotating disc. Around one of the shadows a 4 mm white line was drawn so as to coincide exactly with the penumbra. The diameter of each shadow was approximately 30 cm. The observer sat at a distance of 200 cm behind a large screen with an observation-window (*F*) which could be opened or closed by the experimenter. The observation-window was so placed as to reveal the two color-wheels with their surrounding shadows and their black cardboard backgrounds but conceal the light-source and the shadow-casters. The observation-window could also be fitted, when necessary, with a reduction screen, by means of which the physical equivalence of the two shadows could be tested. 38 observers were used, all drawn from the elementary course in psychology. In terms of previous knowledge of the problem they could all be considered as naïve.

[3] This experiment was performed in 1932 in the Laboratory of Psychology at Cornell University.

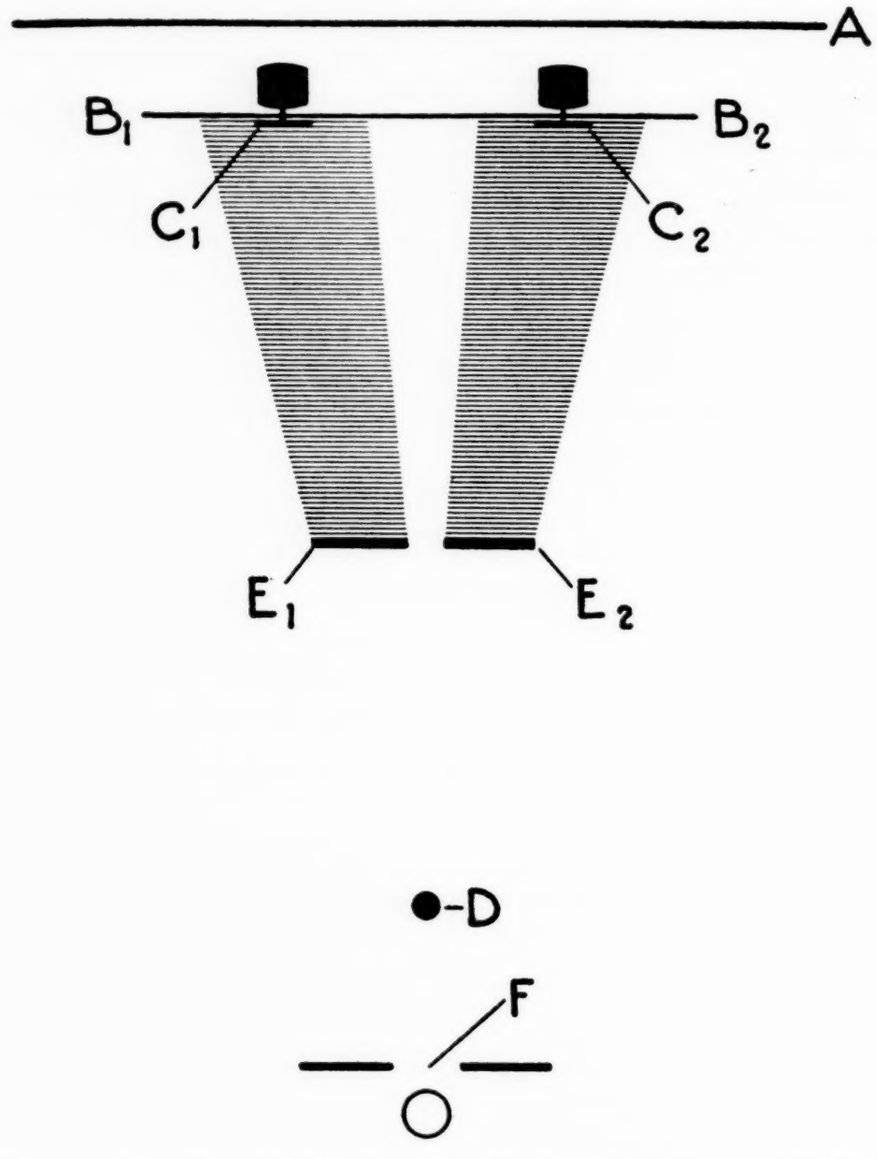


FIG. 1. *A*—uniform black wall; $B_1$, $B_2$—black cardboard screens; $C_1$, $C_2$—color-wheels; *D*—light-source; $E_1$, $E_2$—shadow-casters; *F*—observation window; *O*—observer. Dimensions given in text.

The procedure consisted of three steps, and was uniform for all observers:

*Step I.*—The two color-wheels were fitted with black and white sectors, adjusted in each case at 180° white and 180° black, and set in rotation before the *O* was brought into the room. The light was switched off until the *O* was seated in a chair behind the observation screen, and was then immediately switched on again. After a few minutes for adaptation to the new illumination, the observation-window was opened and the *O* was asked to describe in detail everything he could see, using any terminology, scientific or otherwise, which seemed most appropriate. These descriptions were recorded by the experimenter.

*Step II.*—With the color-wheel in the unencircled shadow remaining unchanged, the color-wheel in the encircled shadow was varied, by means of a modified form of the

method of limits until a brightness[4] equation for the two discs had been obtained. The modification involved merely a shortening of the standard procedure, the equation being based on 4 individual determinations, 2 'up' series and 2 'down' series. It was felt that the resulting loss in precision was less serious than would be the effects of fatigue in a longer series. In any case, the reliability of the brightness differences found is beyond question. No time limit was set for each exposition, but the *O*s were instructed to make their judgments as quickly as possible, shifting fixation back and forth from one disc to the other until a judgment could be formed. The observation time ranged from 3 to 5 secs. for each judgment. *At the conclusion of the equation the Os were asked to examine the apparatus, and the structure of the situation was carefully explained to them.*

*Step III.*—A few days later a new equation was determined under exactly the same conditions as in *Step II*, the only difference being that this time the *O* had full knowledge of the structure of the situation before him.

Table II summarizes the quantitative results of the experiment. The symbols SS, SC and CC indicate the status of

TABLE II

SUMMARIZED RESULTS OF EXPERIMENT WITH UNRECOGNIZED SHADOWS: EXPLANATION IN TEXT

| (1) | (2) | (3) CC | (4) | (5) | (1) | (2) | (3) SS | (4) | (5) |
|---|---|---|---|---|---|---|---|---|---|
| Obs. | Standard | Step 2 | Step 3 | Diff. | Obs. | Standard | Step 2 | Step 3 | Diff. |
| 1 | 180 | 225.0 | 226.3 | 1.3 | 24 | 180 | 243.3 | 236.3 | −7.0 |
| 2 | 180 | 232.5 | 236.2 | 3.7 | 25 | 180 | 237.5 | 216.3 | −21.2 |
| 3 | 180 | 230.0 | 191.3 | −38.7 | 26 | 180 | 192.5 | 202.5 | 10.0 |
| 4 | 180 | 244.3 | 230.0 | −14.3 | 27 | 180 | 212.5 | 212.5 | 0.0 |
| 5 | 180 | 225.0 | 235.0 | 10.0 | 28 | 180 | 227.0 | 227.0 | −17.5 |
| 6 | 180 | 197.5 | 197.5 | 0.0 | 29 | 180 | 231.3 | 231.3 | 7.5 |
| 7 | 180 | 235.0 | 226.3 | −8.7 | 30 | 180 | 217.5 | 217.5 | 2.5 |
| 8 | 180 | 192.5 | 192.5 | 0.0 | | | | | |
| 9 | 180 | 185.0 | 197.5 | 12.5 | Sum | | 1569.6 | 1543.4 | |
| 10 | 180 | 196.3 | 212.5 | 16.2 | M | 180 | 224.2 | 220.5 | |
| 11 | 180 | 206.3 | 208.8 | 2.5 | SD | | 17.8 | 10.8 | |
| 12 | 180 | 195.0 | 217.5 | 22.5 | | | | | |
| 13 | 180 | 218.8 | 208.8 | −10.0 | | | SC | | |
| 14 | 180 | 195.0 | 190.0 | −5.0 | Obs. | Standard | Step 2 | Step 3 | Diff. |
| 15 | 180 | 195.0 | 207.5 | 12.5 | 31 | 180 | 218.8 | 202.5 | −16.3 |
| 16 | 180 | 200.0 | 225.0 | 25.0 | 32 | 180 | 282.5 | 251.3 | −31.2 |
| 17 | 180 | 215.0 | 217.5 | 2.5 | 33 | 180 | 225.8 | 220.0 | −5.8 |
| 18 | 180 | 220.0 | 215.0 | −5.0 | 34 | 180 | 237.5 | 227.5 | −10.0 |
| 19 | 180 | 202.5 | 205.0 | 2.5 | 35 | 180 | 222.5 | 226.3 | 3.8 |
| 20 | 180 | 205.0 | 210.0 | 5.0 | 36 | 180 | 197.5 | 205.0 | 7.5 |
| 21 | 180 | 212.5 | 201.3 | −11.2 | 37 | 180 | 202.5 | 212.5 | 10.0 |
| 22 | 180 | 222.5 | 215.0 | −7.5 | 38 | 180 | 231.3 | 222.5 | −8.8 |
| 23 | 180 | 200.0 | 192.5 | −7.5 | | | | | |
| Sum | | 4850.7 | 4859.0 | | Sum | | 1818.4 | 1767.6 | |
| M | 180 | 210.9 | 211.3 | | M | 180 | 227.3 | 220.9 | |
| SD | | 15.8 | 13.8 | | SD | | 27.1 | 14.4 | |

[4] The meaning of 'brightness' had been established in Step I. If any doubt seemed to exist, it was defined as a 'nearness to white.'

the two shadows as they were apprehended by the *O* in Steps I and II, and are to be interpreted as follows: SS—both shadows apprehended as shadow; SC—unencircled shadow apprehended as shadow, encircled shadow apprehended as color; CC—both shadows apprehended as color. Columns 3 and 4 give the first and second equations respectively, expressed in terms of degrees of white on the color-wheel in the encircled shadow with the other color-wheel constant at 180° W, 180° B. Column 5 gives the differences between the two equations in terms of degrees of white. The results of the three steps may be discussed separately.

*Step I.*—A detailed phenomenological description may be omitted, since the reports of the *O*s did not differ radically from those which we have come to expect in color-constancy experiments. All *O*s found the task of describing what they saw difficult, and were embarrassed by their lack of an adequate vocabulary. Such terms as brightness, lightness, clearness, etc., did not suffice to characterize the differences between the two halves of the field, and frequently such terms as feathery, spongy, soft and fuzzy, were employed. Without exception both the encircled shadow and the disc at its center appeared distinctly darker than the unencircled disc and shadow, and in many cases the encircled disc bore a tinge of brown which was absent from the other. This is, of course, an example of the commonly observed fact that any stray chromatic light, in this case reflected from the slightly brownish floor, may under appropriate conditions become effective as surface-color. In many cases the two shadows produced different impressions of depth, but, apart from the observation, the unencircled shadow was more difficult to localize in a single plane, no strikingly significant generalizations can be made from the reports.

Of the 38 *O*s, 7 reported immediately or after deliberation that what they were observing was two shadows; 8 reported that the unencircled portion was a shadow and the encircled portion a colored surface; and 23 saw both shadows as colored surfaces. In every case the *O* at the end of his description was asked to advance an hypothesis as to how the experi-

mental situation had been constructed, and even then not one of the 23 referred to succeeded in guessing that shadows were being used. The most common hypothesis was that the surfaces had been painted different shades of grey, and the reports frequently included uncomplimentary remarks about the experimenter's ability as a painter. Although the color-wheels were so well-constructed as to produce no perceptible movement in the rotating discs, the sound of the motors was clearly audible, and the suggestion of something rotating was consequently very strong. In several cases it was assumed that the shadows themselves, rather than the discs, were doing the rotating. As soon as Step II was begun, of course, it became clear that the grey surface at the center of each shadow was a color-wheel.

We may infer from these observations that in approximately 60 percent of the cases there was no perception of a special or anomalous illumination. All that the *O*s could say about the illumination was that it came from a concealed source in front of them. Even in those cases in which the *O* described one or both of the shadows as shadow, one may still raise the question as to whether he actually saw it as shadow or simply made a clever inference. It is a common observation in the spot-shadow experiment that in spite of one's knowledge one cannot actually see the encircled shadow as shadow. Fortunately for the thesis here defended, the issue is not a particularly important one. If the *O*s who gave SS reports saw one shadow as shadow and the other as color, which seems most probable, they can be included among those who gave SC reports. If they saw both shadows as shadow, the thesis that the shadow-appearance is irrelevant to the basic differences in brightness becomes thereby stronger.

*Step II.*—In view of the wide dispersion in the quantitative results, the mean equations for the three groups of *O*s cannot be considered as in themselves of significance. According to the results obtained, the mean equations: SS—224.2, SC—227.3, CC—210.9, might suggest that in comparison with the other two groups the CC group showed a smaller brightening effect, and consequently a lower degree

of color-constancy, than did the other two groups. If we group the SS and SC results together, however, and test statistically for the reliability of the difference between their combined mean and that of the CC group, we obtain a standard error of the difference of 7.3, which gives a critical ratio of only 1.7. To obtain a reliable difference between the two groups we should have to use a much larger number of subjects and a more precise method of measurement. Should such a difference be established, the following two hypotheses would have to be tested: (1) that the impression of shadow produces a change of attitude in the *O*, or a loosening of the structure of the field, which facilitates the emergence of color-constancy phenomena; (2) that the ease or difficulty with which a shadow is perceived as a shadow is a manifestation of a fundamental property of the perceptual system of each person, a property which expresses itself correlatively in a high or low degree of color-constancy. The available evidence would suggest that both hypotheses are correct, that they represent, in fact, merely different approaches to the same central problem, namely that of the relation between perceptual attitudes and perceptual phenomena. Unfortunately the question of individual differences in color-constancy has not yet been exhaustively studied.[5]

More important than the differences among the three groups of *O*s, however, are the similarities. Without exception the equations showed that the disc in the unencircled shadow was distinctly brighter than that in the encircled shadow; and this relationship obtained whether the shadows were apprehended as shadow or as color. The equations are not exactly comparable with those obtained in the preliminary experiment, since the standard greys which were used were not the same and the conditions of observations were much more difficult in the main experiment. It is interesting to note, however, that whereas the mean difference between the two sets of equations in the preliminary experiment was 43.7°, the differences between the two situations (*i.e.*, the equation minus 180°) for the SS, SC and CC groups were respectively

[5] A review of the correlational studies is included in Sheehan (14).

44.2, 47.3 and 30.9. There can be no challenging the fact that in the two experiments we are dealing with essentially the same phenomenon.

The following conclusions may accordingly be drawn from the results of Steps I and II: (1) The mere fact that one shadow is encircled and the other is not results in a significant difference in brightness in their respective infields, even when the shadows are not perceived as shadows. (2) The difference is not reliably smaller when neither shadow is perceived as shadow than when one or both are perceived as shadow. (3) This difference is unquestionably a special case of what has traditionally been referred to as color-constancy.

*Step III.*—The results of Step III are difficult to interpret in a positive way. We have no *a priori* grounds for belief that a change in the *O*s understanding of the situation should affect his perceptual attitude to any significant degree.[6] If one were to accept a crudely Helmholtzian position, and attempt to explain color-constancy phenomena as due to unconscious errors in judgment, the present results might serve as further evidence against such a view. The mean differences between the two equations are insignificant in all three groups, and the differences are scattered almost equally in the plus and minus directions (*cf.* Table III). In other words, with a

TABLE III

DIRECTION OF CHANGES REGISTERED IN STEP III; NUMBERS REPRESENT *O*S

| Appearance | + | − | 0 | Total |
|---|---|---|---|---|
| CC | 12 | 9 | 2 | 23 |
| SS | 3 | 3 | 1 | 7 |
| SC | 3 | 5 | 0 | 8 |
| Total | 18 | 17 | 3 | 38 |

fuller understanding of the situation on the part of the *O*, phenomenal equality does not evince any significant change in the direction of physical equality. Fortunately the Helmholtzian argument is nowadays seldom advanced, so that our

[6] *Cf.* Henneman (4), p. 62ff. As one variation in his experiments in brightness-constancy, Henneman compared equations before and after instruction as to the real albedos of the surfaces involved, and found only unreliable differences.

conclusion from Step III simply confirms for the present experiment what has been found by many investigators. We may conclude that the phenomena which we are here studying depend on factors which are independent of our knowledge of the physical structure of the situation.

## Supplementary Experiment on the Effects of Intentionally Changed Attitudes [7]

The results of Step III raise numerous questions about the possible effects on the equations of intentionally changed attitudes in the *O*s. In the following experiment the *O*s were permitted to make a naïve equation first. The situation was then analyzed and explained to them, and immediately afterwards they were instructed to disregard as completely as possible the phenomenal differences between the two shadows, and to make an equation which corresponded to the physical situation. Such tricks as squinting, observing through gaps between the fingers, etc., were, of course, prohibited. The experimental situation was similar to the one already described, with the exception that the observation distance was reduced to 80 cm. The *O* was brought into the room blindfolded, and seated on the floor immediately in front of the cardboard screens. The light and shadow-casters were thus behind the *O*, and consequently had to be raised somewhat so that the *O*s shadow would not fall on the screens. In keeping with Katz's first law of field size (**9**) this position was much more favorable to the production of color-constancy effects than was the one adopted in the previous experiment. The disc in the encircled shadow, set at 360° white, was used as standard, and the equations were made on the disc in the unencircled shadow. The method of limits was used as before, except that in this case 10 determinations (5 'up' series and 5 'down' series) were made for each equation.

Table IV presents the quantitative results of the experiment. Columns 3 and 6 give the brightness equations, expressed in terms of degrees of white, registered on the color-

[7] The following experiment was performed under the writer's direction by Miss Kristel Fuchs in the Psychological Laboratory of Swarthmore College. The *O*s were undergraduate students at Swarthmore.

wheel in the unencircled shadow under the two sets of instructions. Columns 4 and 7 give the corresponding differences between these equations and the standard disc (360° white) in the encircled shadow. Each of these equations represents the mean of 10 separate determinations. The average deviations of each group of 10 are presented in columns 5 and 8.

Three points are worth noting in connection with these results: (1) In spite of the difference in instruction, the two sets of equations do not differ significantly from each other. We

TABLE IV

EFFECT OF INTENTIONALLY CHANGED ATTITUDE

| | | 'Naïve' Attitude | | | 'Physical' Attitude | | |
|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 1 | 360.0 | 171.0 | 189.0 | 12.6 | 174.0 | 186.0 | 17.6 |
| 2 | 360.0 | 161.0 | 199.0 | 12.9 | 156.5 | 203.5 | 28.2 |
| 3 | 360.0 | 144.0 | 216.0 | 3.6 | 144.0 | 216.0 | 22.3 |
| 4 | 360.0 | 140.0 | 220.0 | 8.6 | 142.0 | 218.0 | 16.9 |
| 5 | 360.0 | 122.5 | 237.5 | 18.3 | 126.0 | 234.0 | 14.4 |
| 6 | 360.0 | 120.5 | 239.5 | 9.7 | 104.0 | 256.0 | 11.9 |
| 7 | 360.0 | 119.0 | 241.5 | 7.9 | 115.5 | 244.5 | 26.0 |
| 8 | 360.0 | 119.0 | 241.0 | 9.0 | 110.0 | 250.0 | 22.0 |
| 9 | 360.0 | 118.5 | 241.5 | 10.8 | 112.0 | 248.0 | 18.7 |
| 10 | 360.0 | 108.0 | 252.0 | 7.2 | 112.0 | 248.0 | 19.4 |
| 11 | 360.0 | 108.0 | 252.0 | 7.2 | 108.0 | 252.0 | 18.7 |
| M | 360.0 | 130.1 | 229.9 | 9.8 | 127.6 | 232.4 | 19.6 |
| Sigma | | 20.1 | | | 22.1 | | |

(1)—observer; (2)—standard disc in encircled shadow; (3) and (6)—equations for (2) on disc in unencircled shadow; (4) and (7)—differences between standard and final equation; (5) and (8)—A.D.s of individual equations. All values represent degrees of white on the color-wheel.

may conclude, then, that under these conditions even a deliberate attempt to adopt a 'physical' attitude does not enable the observer to overcome the effects of the differently organized fields. (2) If the average deviation may be accepted as a crude measure of precision, we may conclude that the major effect of the changed instruction was to lower the precision of the judgments. This is supported by the reports both of the *O*s and of the experimenter. At the beginning of the 'physical' series the *O*'s tendency was always to make a deliberate

correction by choosing a value which was obviously too dark, and then in successive determinations to swing back in the direction of the 'naïvely' perceived equality. In many cases, in fact, this swing actually led to an overcompensation in the opposite direction. (3) In spite of the fact that the brightness values in this and in the preceding experiment cannot be reduced to a common absolute scale, it is immediately evident that the brightness differences are even more marked here than they were in the previous experiment. If a less favorable experimental situation had been chosen, the change of instruction might have had a more significant effect.[8]

Until the concept of perceptual attitude has been clarified, and further experiments have been performed, the conclusions drawn from this experiment may be considered as valid only within the limits of the experiment itself. It is clear that the brightness differences here observed are so compelling as to resist the influence of any formal instruction. There can be no doubt, too, that this represents the state of affairs in everyday perception. We cannot, in a fully illuminated room full of dark objects, reverse the relationship between object-color and impression of illumination, and by an act of will see all the objects as bright and the illumination as dark. The structure of the situation is altogether too stable. There are cases, however, as every investigator in this field knows, in which analogous reversals do take place. What appears at one moment as a spot of light in reduced general illumination may suddenly transform itself into a dark color in a strong local illumination; and the alternative appearances may continue to fluctuate back and forth for some time. In such cases the reversal may be facilitated by slight changes in stimulus organization.[9] Sometimes, however, the reversal seems to take place of its own accord, or even in response to the voluntary effort of the

[8] This may account in part for the apparent conflict between these results and those of Henneman (4; p. 63 ff.). Henneman found that by careful training he could induce a reliable shift in perceptual attitude in some of his subjects.

[9] *Cf.* Gelb (2). In an otherwise dark room Gelb focussed a beam of light on a black disc so as to make it coincide exactly with the boundaries of the disc. The disc in consequence looked white. He then introduced a piece of white paper in front of the disc, and the disc immediately ceased to appear as white in the reduced illumination of the room, but was seen as black in the bright special illumination of the beam of light.

observer. In such cases one is tempted to speak of a change in perceptual attitude, *i.e.*, a change within the perceptual system of the observer which brings about corresponding changes in the field. The problem seems to be exactly analogous to that of reversal in spatial designs. Reversals do not take place under ordinary conditions of stable organization; they begin to appear only when there are alternative sets of determinants of approximately equal strength. Although under such conditions a certain amount of voluntary control seems to be possible, the reversal does not appear to be regulated entirely by the effort of the observer. It may take place even when the observer is endeavoring to resist it; and, in any case, the function of the observer's volition seems to be to set the stage for the reversal rather than actually to cause it.

It is clear, then, that the study of these perceptual attitudes in connection with the phenomena of brightness constancy will have to be made under conditions analogous to those which give the best spatial reversals, *i.e.*, in a field in which the determinants of perception are less stably organized than was the case in the foregoing experiment.

## Conclusions

I. To state one major conclusion in a positive way, we may assert that some of the phenomena of color-constancy are independent of the perception of a special or anomalous illumination. This contradicts the common assumption that the constancy of color is maintained by means of a special process whereby the prevailing illumination is independently perceived and 'discounted' or 'corrected for'; and it supports the view that a finally satisfactory explanation of color-constancy will have to be sought in the elementary conditions of field dynamics. Against such a conclusion may be argued: (*a*) that our experiments have not contained valid examples of color-constancy, and (*b*) that our conclusion cannot be applied to all, nor even to the most striking, cases of color-constancy. The first objection may be disposed of simply. If color-constancy phenomena are to be considered, by definition, as only those in which a color is perceived as constant in

spite of an illumination which is perceived as changed or changing, then our crucial cases in the second experiment must obviously be excluded. This leaves us in the ridiculous position of admitting the SC cases as valid but excluding both SS and CC cases, whereas the observations of all three groups were, except for their impressions of illumination, strictly comparable. If, on the other hand, we restrict the term 'color-constancy phenomena' to those cases in which actual changes in physical illumination have been introduced, we should have to reject Hering's spot-shadow experiment and Katz's technique of 'reduction,' both of which have been accepted as valid since the problem was originally formulated. In the common examples of color-constancy we have a phenomenal equality (*i.e.*, equality of brightness) in spite of physical difference (*i.e.*, difference in intensity of reflected light). In our experiments we have held the physical situation constant as to intensity of reflected light and produced a phenomenal inequality. There can be no logical objection to such an inversion of procedure. The phenomena we obtain are still phenomena of 'color-constancy.'

The second objection finds us in a more vulnerable position. We have demonstrated that some color-constancy phenomena are independent of a special impression of illumination, but we cannot prove that the impression of illumination is irrelevant to all such phenomena. The apparent importance of such an impression in everyday perception suggests two possible hypotheses: (1) that there are two causal principals involved, both of which bring about the same changes in brightness, one of which rests on the perception of illumination as such and one of which rests on more obscure conditions of field organization; (2) that there is only one causal system involved, that this is to be described in terms of field-organization, and that the impression of illumination is an epiphenomenon, a parallel but not necessary result of the operation of this system.[10] Of the two hypotheses, the second is the simpler, and for that reason preferable; but until more evidence is available the first cannot be eliminated.

[10] We reject as unworthy the logical alternative that the impression of illumination is always there as a causal factor, but sometimes operates unconsciously.

II. So far, we have accepted without question the term 'color-constancy,' in spite of the fact that it has proved a continual source of embarrassment. The scientific value of the concept of 'phenomenal constancy' has already been questioned by other writers,[11] and other terminology has been proposed. The present study strengthens the objection to the concept, but fails to support some of the alternative formulations.

The principle of phenomenal constancy in its traditional formulation derived its significance from the fact that traditional perceptual theory postulated a primary 'pre-perceptual' correlation between stimulus and sensation. The principle served, consequently, to indicate important discrepancies between physical and phenomenal change. Color, shape and size do not change in exact accordance with changes in retinal stimulation, but seem to be coördinated rather to the relative invariance of objects in the physical world.[12] As such, the principle has been of immense service in the experimental attack on problems of perception, and the 'constancy' phenomena have become basic to the construction of a more adequate perceptual theory. There are several reasons, however, for considering the concept of constancy as a superfluous, if not a dangerous, concept in any scientific treatment of the problems of perception. In the first place, the principle as formulated implies the very hypothesis which it has helped to demolish, namely that of a primary correlation between stimulus and sensation. Having relinquished this hypothesis, it is misleading for us to continue to talk in terms of discrepancies between the physical and the phenomenal. Such discrepancies no longer constitute a problem. Secondly, our use of the term tends to result in the singling out of certain groups of phenomena which can be so designated and considering them as special cases, governed by special laws. Gelb argued in 1929 (**2**), and all subsequent writers have agreed with him, that there is nothing unique about the phenomena of color-constancy. The 'normal' color of Katz is in no sense an

[11] *Cf.* Gelb (**2**), Thouless (**15**), Helson (**3**).

[12] *Cf.* Koffka's distinction between proximal and distant stimulation (**12**, ch. 3).

absolute, since each situation creates its own standard of normality, and we are consequently not justified in speaking of the 'central transformation' of a color impression in the direction of the normal. The perception of a color in an 'abnormal' illumination is no more, and no less, a problem than is the perception of a color in a 'normal' illumination. The central problem is, rather, to explain why any given color appears as it does under any given set of conditions.[13]

A third, and much more serious, difficulty with the concept of constancy is that it is essentially normative in character. Without disparaging the value of a functional approach to the study of perception, we must recognize that the postulate of a 'tendency in the organism to hold perceptual objects constant' runs the danger of becoming an explanatory principle. There can be no questioning the fact that under ordinary conditions of living such a result is actually achieved, and that if it were not, the organism would have no stable world within which to live. From an evolutionary point of view, too, we may assume that the equipment responsible for such an achievement has developed because it is of survival value. We might even be justified in linking it with the processes which underlie homeostasis, and considering both as revealing a general biological principle. In the present state of our knowledge, however, it is dangerous to invoke such a general principle when we are attempting to give a causal account of specific phenomena, for it leads us to assign arbitrary value to a phenomenon in accordance with its degree of approximation to the norm.

Recognizing this difficulty, Thouless (**15**) and Brunswik (**1**) have independently contended that the phenomenal object should be considered as an intermediate, or compromise between two poles, at one of which is the stimulus complex, and at the other the 'real' (physical) object, a point of view which Woodworth has recently supported (**17**, ch. 24). Thouless has

[13] Helson (**3**) has admirably defended this thesis in his study of the color of non-selective samples in chromatic illumination. It is to be hoped that further experimentation will enable him to extend his principle to include changes in color related directly to changes in field-organization.

suggested "phenomenal regression to the 'real' object" as an alternative to 'phenomenal constancy.' Aside from its unintentional implication of a temporal course, which exposes it to the attack directed against the concept of transformation, the concept of phenomenal regression may also be criticized as being normative; it postulates a norm in the 'real' world which the phenomenal object approaches, and in terms of which it is evaluated. This, too, is a position which might conceivably be defended on epistemological grounds, but which should come at the end and not at the beginning of a psychological inquiry.

Our final objection to the concept of constancy is that it cannot be applied meaningfully to such observations as we have been reporting; yet, as we have already pointed out, these phenomena are essentially the same as those which conform more exactly to the traditional definition of color-constancy. In such an artificial situation as we have described, which gives us brightness differences which are referable solely to differences in the organization of the field, and which do not involve the intervention of special impressions of illumination, there is no suggestion of a 'constancy of brightness.' We might, of course, consider them as curiosities or aberrations of a color-constancy mechanism, analogous to the 'illusions' of space-perception. That, however, would be to reject the possibility that the determinants of brightness, like the determinants of space, can be coördinated into a unitary system of laws. Nor does the concept of phenomenal regression prove to be any more serviceable. Neither of the aforementioned shadows can be said to be more 'real' than the other. Unless we stretch both our logic and our epistemology, we cannot consider either degree of brightness as a compromise between two end states. Each degree of brightness is valid for the conditions under which it is perceived, and if we are to seek an explanation, we must do so by means of a more precise formulation of those conditions. It is not likely that the term 'phenomenal constancy' will be dropped from our psychological vocabulary. It represents an important chap-

ter in the history of experimental psychology, and it can still assist in the initial formulation of new problems. It seems highly desirable, however, that further investigations of 'constancy' problems should be conducted within a different framework.

III. On the positive side, the present experiments suggest the possibility that at least some of the phenomena of color-constancy may be incorporated within a general theory of color-perception which gives due weight to conditions in the surrounding field. The fact that brightness differences may be induced by means of changes in field organization without any accompanying changes in stimulus intensity points toward the existence of an important correlation between brightness and certain factors of field organization.[14]

The crucial factor which seems to be operative in the present experiments is the contour. The substitution of a precise contour for the typical blurred contour of the shadow results in a decrease in the brightness of the included area, *i.e.*, in an enhancement of the brightness difference between infield and circumfield. In the somewhat similar *Florkontrast* experiment, which also involves blurring of contours, exactly the opposite result is obtained. The greatest mutual enhancement of brightness differences is found when the fields are seen through a blurring medium. This apparent contradiction must obviously be resolved by further experimentation, and the logical point of attack lies in a further study of the functions of the contour. In a later report the author will describe experiments in which contours were artificially blurred and will examine the implications of these for a theory of brightness-perception.

[14] Such a correlation has already been worked out in detail for the standard phenomena of contrast (16). Although a restatement of the problem of visual contrast is sorely needed, it is clear that the differences in brightness here reported cannot be explained in terms of contrast in any of its traditional meanings. It seems more likely that the laws of surface contrast will prove to be special cases of more general laws which govern the interaction of areas in the visual field.

## Summary

When the penumbra of a shadow is destroyed by an encircling line, the region within the line becomes darker, a phenomenon which is usually accepted as an example of color-constancy. In the present experiment two equivalent shadows, one encircled and the other unencircled, were used, and measurements of brightness difference were made by means of rotating discs inserted within the shadows. It was found: (1) that the characteristic difference in brightness was present even when neither shadow was perceived as shadow; (2) that this difference persisted even after full explanation of the experimental conditions to the *O*s; (3) that the difference persisted even when the *O*s were instructed to disregard the phenomenal difference in brightness and equate for the physical properties of the situation. It is argued: (1) that the impression of anomalous illumination is not a necessary condition of some color-constancy phenomena; (2) that some of these phenomena are more properly referable to conditions of field-organization; (3) that the concept of 'constancy' is now of dubious value in the study of the determinants of color.


(Manuscript received January 14, 1940)

## References


1. Brunswik, E., *Wahrnehmung und Gegenstandswelt*, Leipzig, 1934. Pp. 244.
2. Gelb, A., Die 'Farbenkonstanz' der Sehdinge, *Handb. d. norm. u. pathol. Physiol.*, 1929, **12**, 594–678.
3. Helson, H., Fundamental problems in color vision. I. The principle governing changes in hue, saturation and lightness of non-selective samples in chromatic illumination, *J. exper. Psychol.*, 1938, **23**, 439–476.
4. Henneman, R. H., A photometric study of the perception of object color, *Arch. Psychol.*, 1935, **179**, pp. 88.
5. Hering, E., *Grundzüge der Lehre vom Lichtsinn*, Berlin, 1920.
6. Kaila, E., Gegenstandsfarbe und Beleuchtung, *Psychol. Forsch.*, 1923, **3**, 18–59.
7. Katona, G., Zur Analyse der Helligkeitskonstanz, *Psychol. Forsch.*, 1929, **12**, 94–126.
8. Kardos, L., *Ding und Schatten*, *Zsch. f. Psychol.*, 1934, Ergänzungsband 23. Pp. 184.
9. Katz, D., *Der Aufbau der Farbwelt*, Leipzig, 1930. Pp. 484. Eng. trans., *The World of Colour*, London, 1935. Pp. 296.
10. Koffka, K., A new theory of brightness constancy; a contribution to a general theory of vision, *The Physical and Optical Societies. Report of a Joint Discussion of Vision*, Cambridge, 1932, 182–188.

11. ——, Some remarks on the theory of colour constancy, *Psychol. Forsch.*, 1932, **16**, 329–354.
12. ——, *Principles of Gestalt Psychology*, New York, 1935. Pp. 685.
13. MacLeod, R. B., An experimental investigation of brightness constancy, *Arch. Psychol.*, 1932, **135**, pp. 102.
14. Sheehan, M. R., A study of individual consistency in phenomenal constancy, *Arch. Psychol.*, 1938, **222**, pp. 95.
15. Thouless, R. H., Phenomenal regression to the 'real' object, I, *Brit. J. Psychol.*, 1931, **21**, 339–359; II, *ibid.*, 1931, **22**, 1–30.
16. Tschermak, A., Licht und Farbensinn, *Handb. d. norm. u. pathol. Physiol.*, Berlin, 1929, 12 (1), 295–501.
17. Woodworth, R. S., *Experimental Psychology*, New York, 1938. Pp. 889.

# PITCH AND FREQUENCY MODULATION

BY DON LEWIS, MILTON COWAN, AND GRANT FAIRBANKS

*State University of Iowa*

This report is the first of two to be concerned with an investigation undertaken to determine the principal psychophysical relationships which exist between the perception of pitch and the extent and rate of frequency modulation in short pulses of sound. The results of the investigation should have more than theoretical interest: in addition to their obvious bearing on auditory theory, they should relate basically to various aspects of pitch perception in speech and music, and especially in speech. In fact, the investigation is an outgrowth of the interest of the authors in the perception of pitch patterns in speech. Any one who is at all familiar with the physical analyses that have been made of continuous discourse realizes that the fundamental frequency of speech sounds is almost never constant, even for very short intervals of time, but is almost always changing in one way or another. Teachers of speech refer to voice inflections of various types. These inflections arise from changes in frequency (modulation); and both the extent and rate of change are often very great. The mean rate of change, as found by Fairbanks and his students (**2**), is around 17 tones per second in the voices of superior speakers, and changes in excess of two octaves in a tenth of a second have been measured by Cowan (**1**).

At the outset of the investigation, three specific aims were formulated. They were: (*a*) to determine the perceived extent of pitch change as a function of the extent and rate of frequency modulation; (*b*) to determine threshold values for the detection of direction of pitch change, the thresholds to be expressed in terms of either extent or rate of modulation; and (*c*) to evaluate the experimental results in terms of current auditory theory. The present report deals with results related to aim (*a*). Other features of the more general problem

have been investigated by Lewis and Pepinsky and will be treated in a subsequent paper.

The typical stimulus used in the investigation was a short pulse of sound of given duration and given extent and form of frequency modulation. All of the various pulses were produced by modulating the output of an oscillator. Eventually, in order to give an adequate picture of pitch perception in speech, it will no doubt be both necessary and desirable to deal directly with typical speech sounds, as Root (5) and Peters (4) have already attempted to do. However, as a beginning, it was deemed advisable to use stimuli which could be controlled with considerable exactness.

The apparatus employed in producing the stimulus sounds consisted essentially of a rotary-type air condenser which was connected in parallel with the tuning condenser of the fixed oscillating unit of a beat-frequency oscillator. A schematic drawing of the condenser is shown in Fig. 1. As the rotor was revolved, changes in capacitance oc-

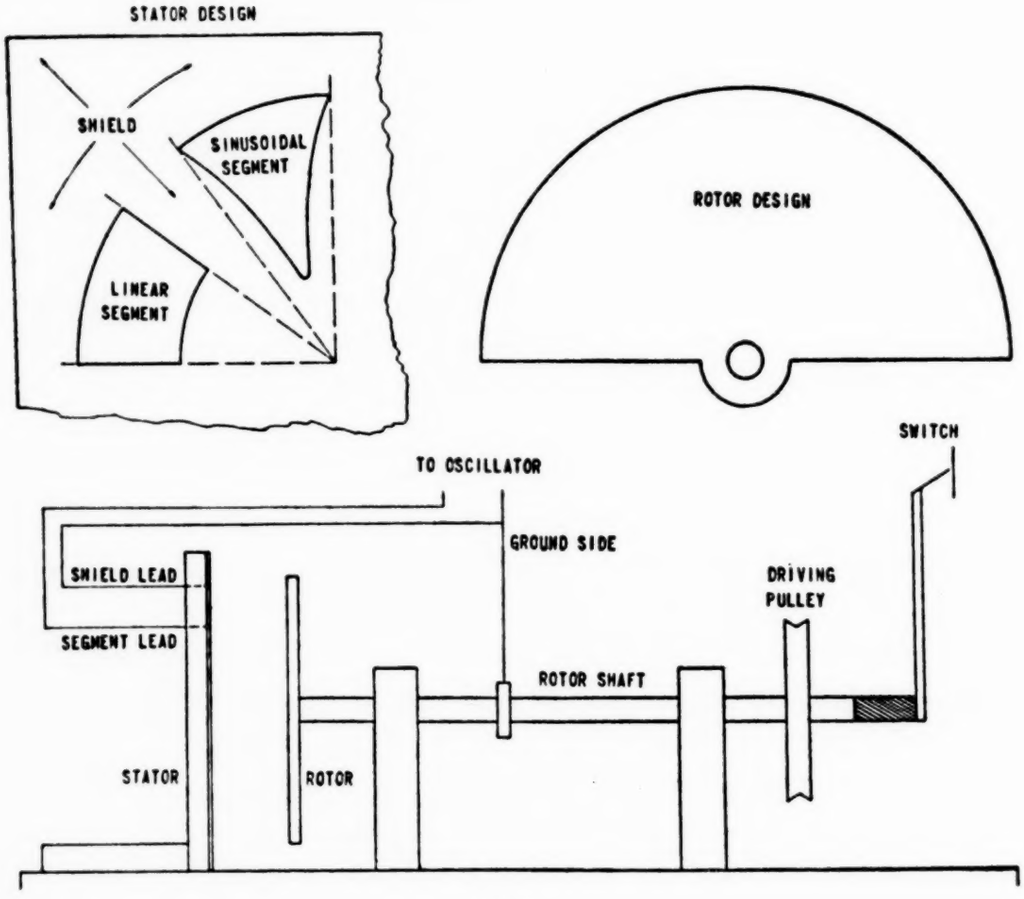


FIG. 1

curred, and these changes induced changes in frequency. Similar instruments have often been used to produce 'warble tones' for reverberation time measurements, and Shower and Biddulph (6) used a similar condenser to produce the stimuli for their investigation of differential thresholds for pitch discrimination. The present apparatus, however, was unique in its provision for electrostatic shielding. As shown in Fig. 1, the rotor was semi-circular in form. When it was either completely over or completely

off the face of the stator, there was no detectable change in capacitance even though the rotor was moving. It was found, however, before the shielding feature was introduced, that capacitance changes occurred as the edge of the rotor approached the edge of the stator. This trouble was eliminated by means of shielding. The construction was as follows: A large square plate was covered over its entire face with aluminum foil, the foil being insulated from the plate by a layer of cardboard. The stator segment proper was then cut out in the foil. All of the foil remained in place, but the stator segment was isolated from the surrounding foil by a narrow air space. The surrounding foil constituted the shielding. It and the rotor shaft, together with the rotor plate, were connected to the ground side of the oscillator. The shielding minimized the 'edge-effect' to a point where a change in capacitance could not be detected until the rotor plate actually began to pass over the face of the stator.

The form of the modulation was determined by the design of the stator segment. Two forms were used—sinusoidal and linear. Designing the linear stator presented no problem. The sinusoidal stator had to be corrected for curvature, and its final design was based upon a method of correction proposed by Kurtz and Larsen (3). When the linear stator was used, the output of the oscillator varied in frequency with time in a manner depicted in the upper part of Fig. 2. The type of frequency change which was

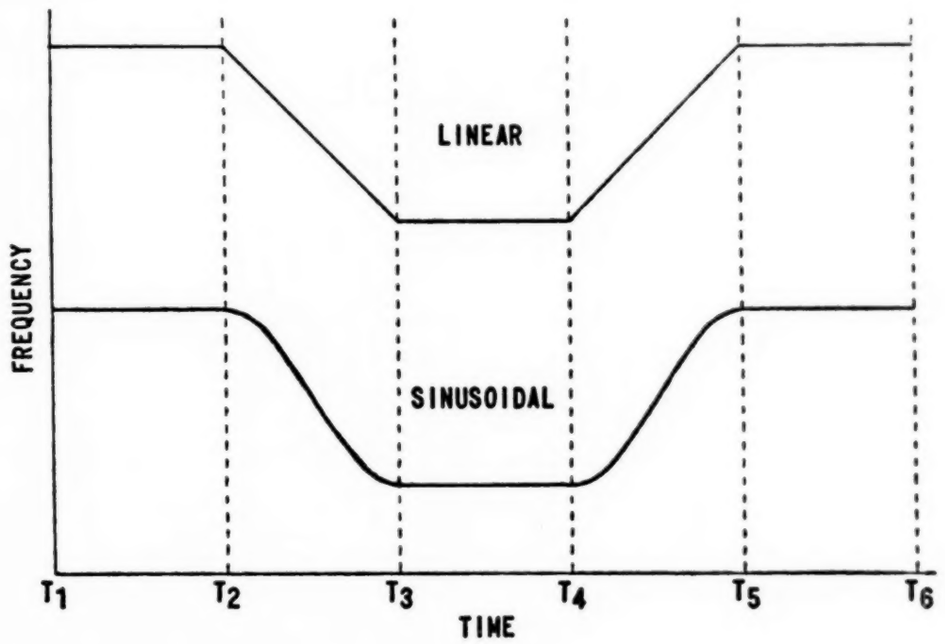


FIG. 2

secured with the sinusoidal stator is shown in the lower part of the same figure. It will be seen that, in either case, frequency remained constant for a certain interval of time and then changed either downward or upward to another steady-state period.

The actual stimuli all were short pulses in which frequency varied from beginning to end. These pulses were secured through the use of a special switch, fastened to, but insulated from, the rotor shaft, which made it possible to 'pick out' any desired portion of the oscillator output. For example, when downward modulation was wanted, the switch was set to pick out only that portion of the oscillator output which lies, in Fig. 2, between $T_2$ and $T_3$. Isolation of the portion between $T_4$ and $T_5$ gave upward modulation. The extent of modulation was controlled by adjusting the distance between the

rotor and stator plates. The duration of a pulse, which naturally determined the rate of modulation for any given extent of modulation, depended upon the rotational speed of the rotor.

The output of the oscillator could be led either into a loud speaker or into earphones; or, as was more usual, it could be led into the cutting head of a phonograph recorder. The stimuli were typically presented to the observers from phonograph records. The phonograph technique was used because of the difficulties involved in changing the condenser from a setting for one rate or extent of modulation to another setting. The various stimuli could easily be recorded separately, each one segregated from all the others, and then could be presented singly. Great care was taken in the recording as well as in the play-back, and the frequency-response characteristics of the recording and reproducing systems were properly equalized over the range of frequencies used. It is believed, therefore, that all important factors were sufficiently well controlled.

Figure 3 shows an oscillogram of a typical stimulus, as reproduced phonographi-

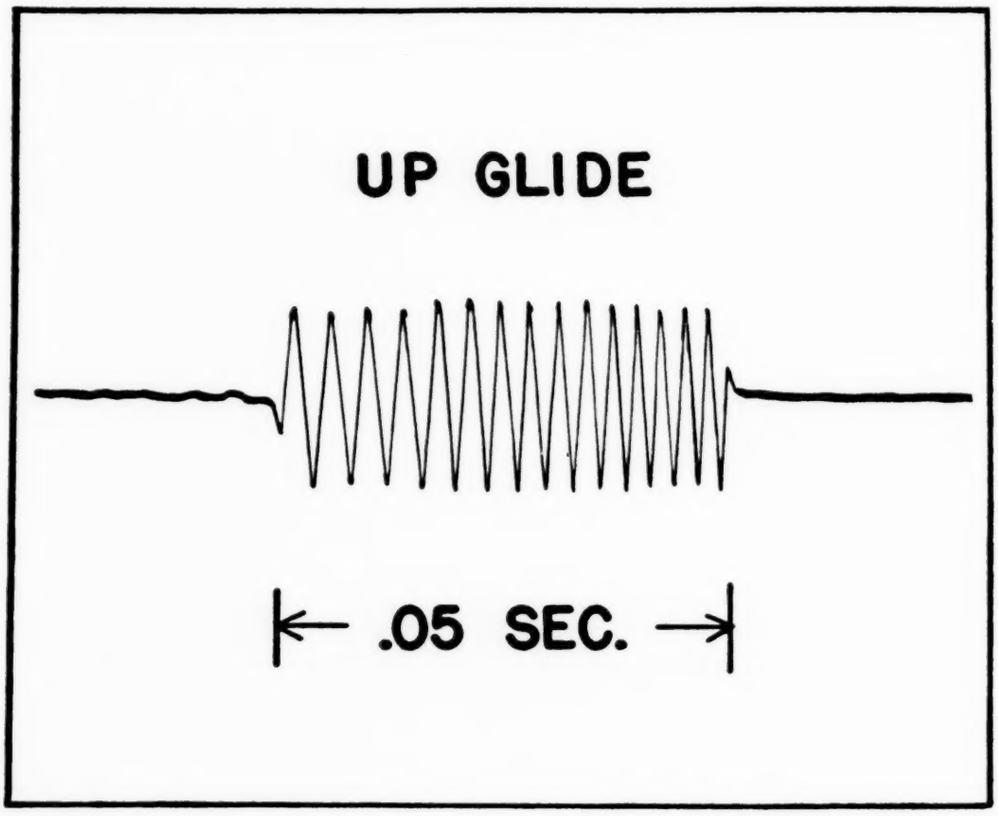


FIG. 3

cally. In this case, the duration of the pulse was 0.05 sec. The modulation was upward and its extent was five full tones. The frequency change was from $A_3$ (= 220 cycles) to $G_4$ (= 392 cycles) with $D_4$ (= 293.7 cycles) as the geometric mid-point of the sweep. As indicated by the small variations in the amplitude of the oscillations in Fig. 3, the intensity of the pulse may not have been absolutely constant but, in any event, it did not vary by a significant amount. All of the phonographically reproduced pulses were without doubt essentially constant in intensity throughout their course.

That the form of the modulation agreed with what was expected from the design of the condenser plates may be seen from the graphs in Fig. 4. The graph at the left presents information on a sinusoidal pulse which had a duration of 0.3 sec. and a downward frequency change from $G_4$ to $A_3$. An oscillogram was first made of the pulse. Measurements were then made on the film of the distances between successive points where the oscillographic tracing crossed the axis. These lengths were converted into temporal values, and the reciprocals of the latter were taken to be indicative of frequency. The dots in the graph depict the measured values, while the curve shows the expected form of the frequency change. The dots all fall very close to the curve. Comparable information on a linear pulse is given in the right-hand graph. Measurements which were made on other pulses revealed that they also had the expected form of modulation.

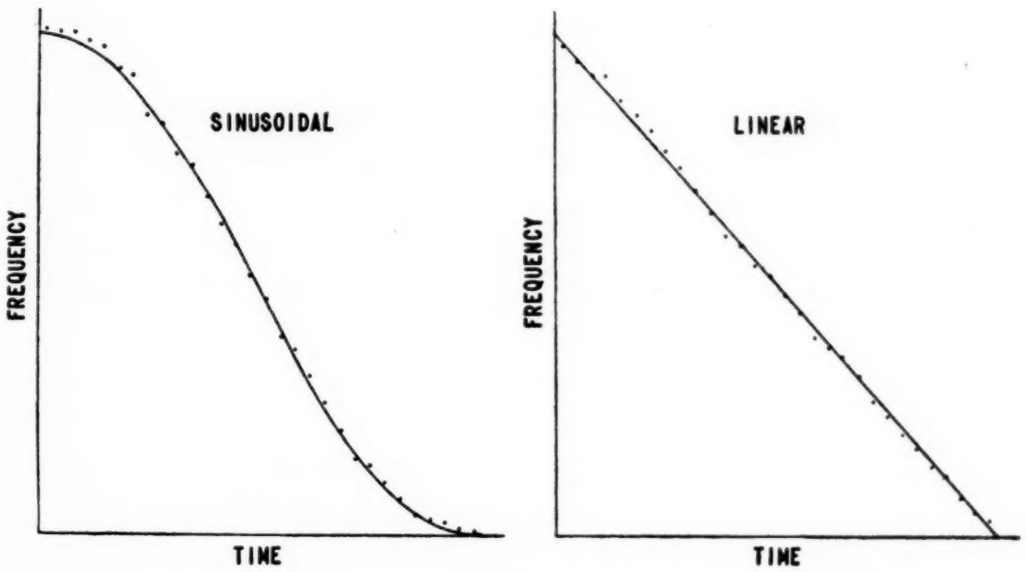


FIG. 4

The first experimental procedure used in attempting to measure the perceived extent of pitch change arising from a stimulus pulse was to present the pulse and then, immediately thereafter, to present one of several different pairs of pure tones, each pair delimiting a certain pitch interval. The observers were asked to judge whether the extent of pitch change in the pulse was greater or less than the extent of the interval. The pulse, whose duration and extent of modulation were known, was always presented first. A given pair of tones, presented successively, followed the pulse. The two tones in a pair each had a duration of 0.5 sec. and were separated by 0.2 sec. The extent of pitch change in the pulse was judged as either greater or less than the pitch extent set off by the succeeding pair of tones.

A typical experimental series involved the use of three sinusoidal pulses which differed in duration but which were alike in the extent and direction of frequency change. The three durations were 0.10, 0.15, and 0.20 sec. The frequency change was downward and extended three tones—from $F_4$ to $B_3$. The geometric mid-point was $D_4$; that is, D above middle C. Five different intervals were used for comparison purposes. Their extents were 1.0, 2.0, 3.0, 4.0, and 5.0 tones ($D\sharp_4$ to $C\sharp_4$, $E_4$ to $C_4$, $F_4$ to $B_3$, $F\sharp_4$ to $A\sharp_3$, and $G_4$ to $A_3$, with the geometric mid-point always $D_4$). For this series, 19 observers were used, all of whom were either graduate students or advanced undergraduate students in psychology. Each of the three pulses was paired with each of the five intervals, and each pair was presented ten times. Consequently, each observer made ten judgments on each pulse in relation to each of the five intervals. The results are presented graphically in Fig. 5. In the figure, percentage of greater judgments

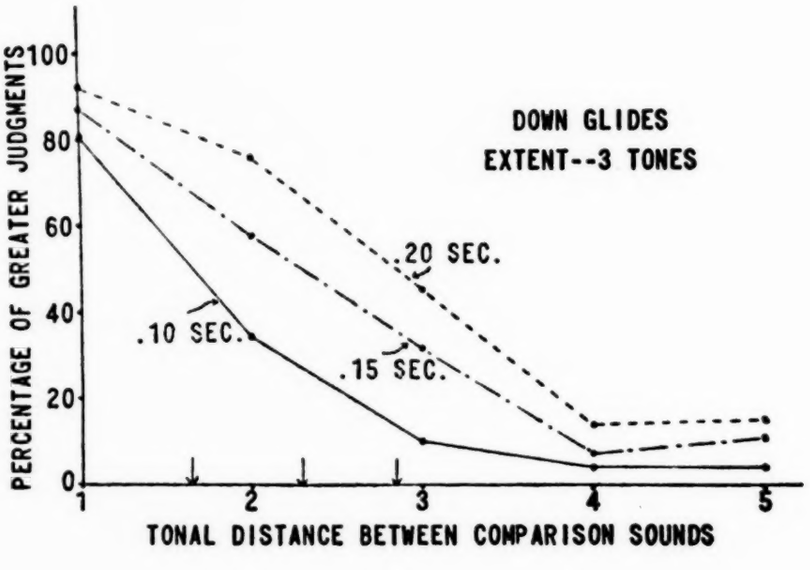


FIG. 5

is plotted against the extent of the comparison intervals. There is a separate curve for each of the three pulses used. Each dot is based on a total of 190 judgments and shows the percentage of times that a pulse of designated duration was heard as having an extent of pitch change greater than the corresponding interval. For example, the pulse with a dura-

tion of 0.15 sec. was judged, in 58 percent of the trials, to have a pitch change greater than an interval of 2.0 tones. The same pulse was judged to have a pitch change greater than an interval of 3.0 tones in only 34 percent of the trials.

The mathematical procedure used in determining the point of subjective equality (P.S.E.) was to compute, by means of linear interpolation, the interval extent which corresponded to a percentage of 50. The three arrows along the abscissa indicate, in order from left to right, the location of the computed P.S.E.'s for durations of 0.10, 0.15, and 0.20 sec. The values were 1.7, 2.3, and 2.8 tones. These values are interpreted to mean that the three pulses were heard as having corresponding amounts of pitch change. Inasmuch as all of the pulses were modulated to the same extent (3.0 tones), extent of pitch change is shown to depend upon duration (or, more exactly, upon rate of modulation).

Additional experimental series of a similar type were run, in which pulses of other durations and other extents of frequency modulation were employed; and all of the results indicated that the perceived extent of pitch change in short pulses of modulated sound is dependent upon both the rate and the extent of modulation.

After exploratory work of this kind had been done, it was decided that a different procedure was perhaps equally, if not more, satisfactory. This procedure was based upon the method of equal-appearing intervals. It involved having trained musicians make direct estimations of the extent of pitch change in various pulses in terms of standard musical intervals and sub-divisions of them. The trained musician typically is able to recognize standard musical intervals (as played on a piano, for example), and he is usually quite adept at detecting when an interval deviates noticeably from some standard extent. For example, if he is asked to judge the extent of an interval which is close to but somewhat less than a perfect 5th, he can ordinarily perceive not only that it is close to the 5th, but that it is closer to a 5th than to an augmented 4th. Or, if an interval happens to fall some place between a major and a minor 3d, he can give not only its

general location (as being between the two) but also a fair approximation of its more exact location (about mid-way between the two, for example).

The musicians who were chosen for observers were instructed to estimate the extent of pitch change which occurred when various pulses were presented to them. By way of making the estimations as uniform as possible, they were asked to conceive of an interval continuum ranging from no extent (unison) up to an extent of a full octave; then, to arrange the various standard musical intervals in appropriate order along this subjective continuum; and finally, to divide each semi-tone division of the continuum into sub-parts. As an example of subdivision, they were asked to think of the minor 3d as having a definite location in the continuum, with the major 3d a proper distance above it. Now, if a pulse gave rise to a perceived extent of pitch change which was judged, on first approximation, to be as great as the general extent of the thirds, they were to make, on more careful scrutiny, one of six possible estimations of extent: 'minus minor 3d,' 'minor 3d,' 'plus minor 3d,' 'minus major 3d,' 'major 3d,' or 'plus major 3d.' This principle of subdivision was applied to the whole continuum, from unison to the octave; so there were, in all, 37 possible estimations of extent. The various divisions of the continuum were to represent equal-appearing extents.

Any attempt to divide a psychological continuum of magnitude or extent into more than 12 or 15 equal-appearing steps is usually regarded as being extremely difficult. For this reason, the use of 37 steps in this investigation may seem quite unreasonable. It should be remembered, however, that trained musicians are able to use the standard musical intervals, which they have learned to identify quite exactly, as reference points (or sign-posts) by means of which they can keep their bearings. All of the musicians who served as observers were quite confident of their ability to make the required estimations. This confidence, together with the fact that the estimations were reasonably consistent from trial to trial, tended to establish the validity of the general procedure.

The practical advantage of the estimation technique over the comparison technique was that it drastically reduced the number of necessary judgments. Each estimation could be used directly in computations of perceived extent. Comparison judgments, on the other hand, could be used only indirectly, and many of them never contributed at all to the final results.

A typical experimental series, in which the estimation technique was employed, will now be described. Forty pulses, all with a duration of 0.10 sec., were the stimuli. Of the 40, 20 were sinusoidal and 20 linear in form. The frequency modulation in half of the pulses of each form was upward in direction; in the other half, downward in direction. In all, then, there were four kinds of pulses: upward-sinusoidal, downward-sinusoidal, upward-linear, and downward-linear. The ten pulses of each kind differed in extent of modulation. The ten extents were 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0 tones. The geometric mid-point of each frequency sweep was $D_4$ (=293.7 cycles).

As indicated above, duration was held constant in this particular series. It was desired to determine, for the specified duration, the essential relationship between perceived extent of pitch change and extent of modulation, and to discover whether the relationship varied with either the form or the direction of modulation. Fourteen musicians served as observers. They were selected by a staff member of the Department of Music from among graduate and undergraduate students in the department, on the basis of their recognized ability to identify and produce standard musical intervals. Most of them were violinists. They were asked to estimate the extent of pitch change in the various pulses, after the manner described above. As a routine procedure, each observer heard three repetitions of a given pulse before making an estimation. This number of repetitions usually sufficed, although a fourth presentation was occasionally requested and given.

In the experimental trials, the stimulus pulses were presented at random. Each observer wrote down his own estima-

tions of the extent of pitch change. Two of the 14 observers each made ten judgments on each of the 40 stimuli, and the other 12 each made two judgments on each of the 40. Thus, there was a total of 44 estimations on each pulse. Means of the various distributions of estimations were computed.[1] Representative SD's also were obtained. The principal results are summarized in Figs. 6 and 7. In both figures, perceived extent in tones is plotted against physical extent in tones. The diagonal line in each figure indicates where the experimental values would have fallen if there were a one-to-one correspondence between perceived and physical extents. Figure 6 gives information on the two kinds of linear pulses

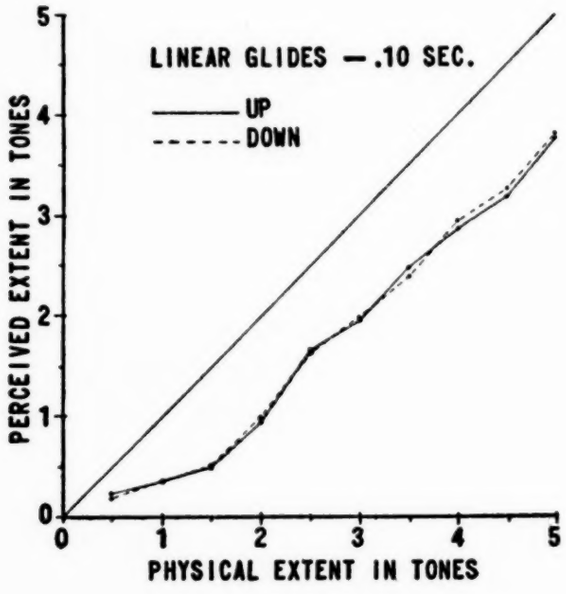


FIG. 6

The dots, which have been connected by straight lines, give the means of estimations of extent for pulses with the physical extents indicated along the abscissa. When the extent of modulation was greater than 1.0 tone, there tended to be almost a straight-line relationship between perceived and physical extents. In other words, for physical extents greater than 1.0 tone, increases in perceived extent tended to accom-

[1] The procedure was to assign each estimation of extent an appropriate numerical equivalent. The numerical values were used in computing means and SD's. Then, the computed means and SD's were converted into tonal extents.

pany equivalent increases in physical extent, although perceived extent lagged behind physical extent by something close to 0.75 tone.

Inasmuch as the two curves in Fig. 6 practically coincide, it may be concluded that, for the group of linear pulses investigated, the direction of modulation did not have a significant bearing on perceived extent. A similar conclusion on direction of modulation holds for the sinusoidal pulses investigated. In fact, no evidence was found in the entire investigation to indicate that direction of modulation has any significant influence on perceived extent. For this reason, in the preparation of Fig. 7, the results for the two kinds of linear

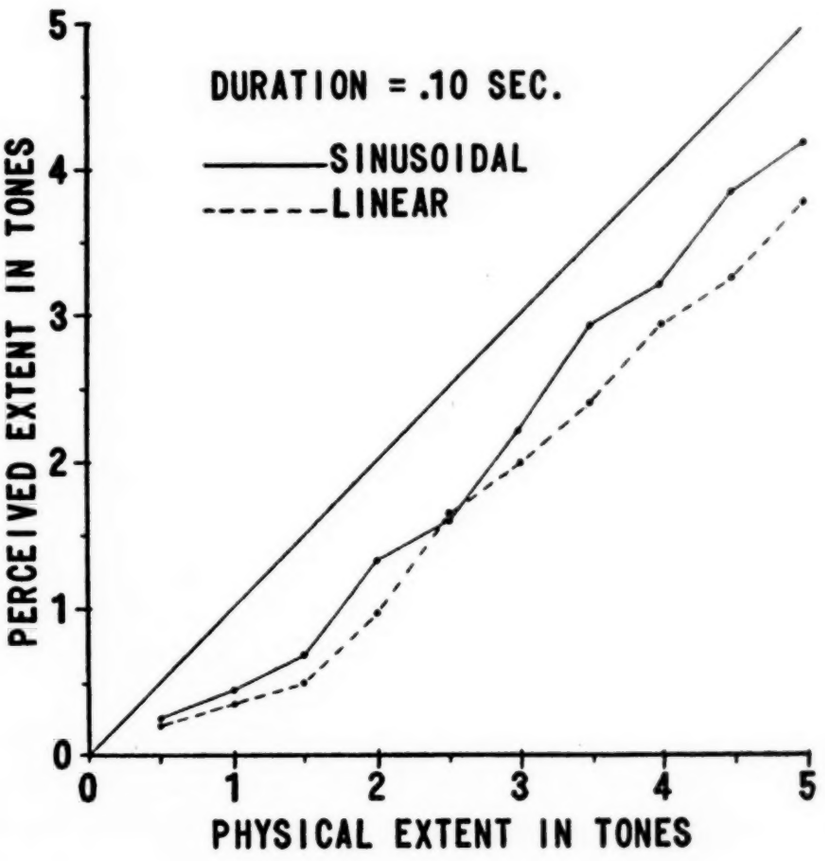


FIG. 7

pulses were combined, as were the results for the two kinds of sinusoidal pulses. In Fig. 7, therefore, the curves portray the perceived extents of sinusoidal and linear pulses, regardless of direction of frequency modulation. As the curves indicate,

the perceived extent of pitch change in sinusoidal pulses was somewhat greater, on the whole, than was the perceived extent of linear pulses. Perceived extent is thus shown to be related to the form of modulation.

The curves in Figs. 6 and 7 are based upon the means of distributions of estimations. That the means could be regarded as reasonably correct was shown by the fact that the SD's of all of the distributions were fairly small. An occasional SD was as large as 0.6 tone or as small as 0.1 tone, but most of them fell between 0.3 and 0.5 tone. The larger SD's were for distributions of estimations on the pulses with the larger physical extents of modulation.

Because the stimulus pulses were presented phonographically through a loud speaker in a reverberant room, the question arose as to whether or not reverberation was a factor in determining perceived extents. As a means of answering the question, one observer who had judged a group of pulses when they were presented through the loud speaker, judged the same pulses when they were presented phonographically through earphones. He made ten observations on each of ten pulses under each of the two conditions of listening. The means of his estimations of extent were not identical, of course, for the two conditions of listening, but they were sufficiently similar to warrant the conclusion that the effects of reverberation were probably negligible.

Additional information on perceived extent is presented in Fig. 8. In this figure, perceived extent is plotted against duration. The data are for a series of sinusoidal pulses in which the frequency modulation was downward in direction. There are five curves in the graph, one for each of five different extents of modulation. The numerals adjacent to the curves indicate the respective extents in tones. Each dot represents a mean of the distribution of estimations of extent of pitch change for a pulse of the designated duration and extent of modulation. The various means were not based upon the same number of estimations of extent. Consequently, they are probably not equally dependable. However, the curves are believed to indicate the basic relationship.

The curves in Fig. 8 show that perceived extent varies as a function of both the extent and rate of modulation. With an increase in the duration of a pulse of given extent of modulation, there was an increase in perceived extent. This fact implies that perceived extent increases with a decrease in rate of modulation. Consider the curve for pulses having an extent of modulation of 5.0 tones. There was an increase in perceived extent of pitch change from 3.3 to 4.4 tones as the duration was increased from 0.025 to 0.30 sec. (or, in terms of

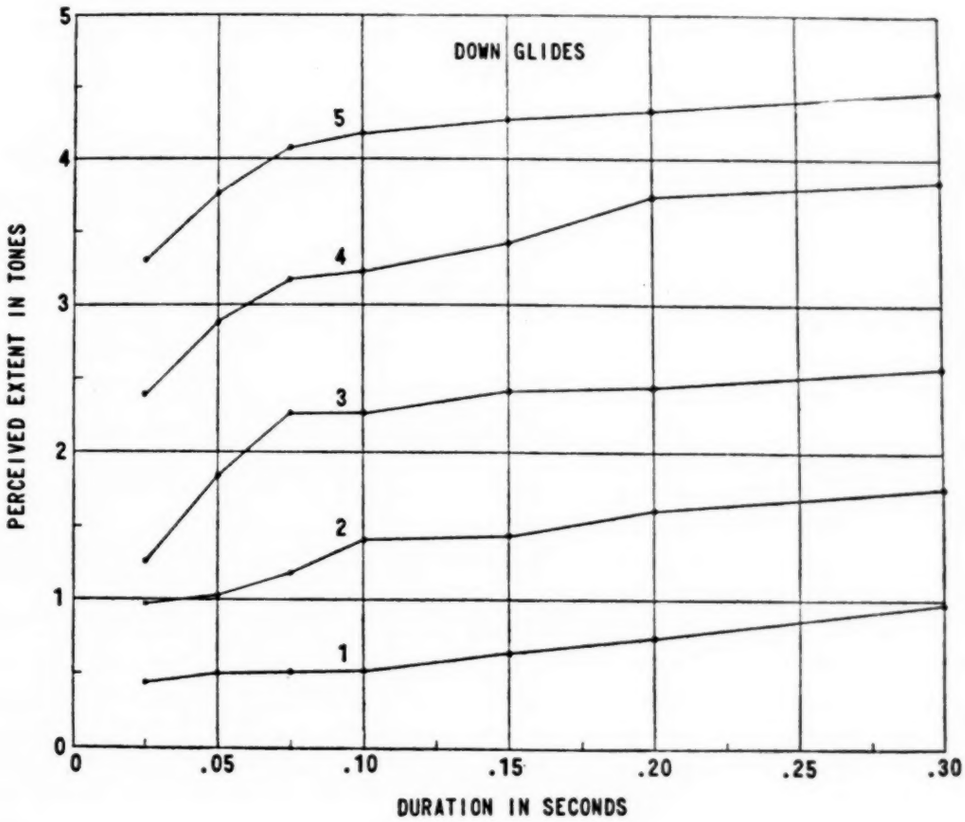


FIG. 8

rate of modulation, as rate decreased from 200 to 16.7 tones per second). Similarly, there was an increase in perceived extent from about 2.4 to about 3.8 tones when the pulses of 4.0 tones in physical extent were increased in duration from .025 to .30 sec. (or when they were decreased in rate of modulation from 160 to 13.3 tones per second). Comparable information can be had for other extents of modulation from an inspection of the other curves in the figure.

As already indicated, the results presented in Fig. 8 are for pulses in which the direction of modulation was downward. Information has also been obtained on a series of pulses in which the direction of modulation was upward. This information shows that, in the case of upward modulation also, perceived extent of pitch change varies with both rate and extent of modulation.

As a whole, the results of the investigation showed that the extent of pitch change arising from short pulses of sound in which frequency is modulated (*a*) increases with increases in the extent of modulation; (*b*) decreases with increases in the rate of modulation; (*c*) is somewhat greater for sinusoidal than for linear forms of modulation of the same extent and rate; and (*d*) is apparently independent of the direction of modulation. The theoretical implications of the findings will be presented, together with data on threshold values for the detection of direction of pitch change, in a subsequent paper by Lewis and Pepinsky.

REFERENCES

1. COWAN, M., Pitch and intensity characteristics of stage speech, *Arch. Speech*, 1936 Supplement.
2. FAIRBANKS, G., Recent experimental investigations of vocal pitch in speech, *J. acoust. Soc. Amer.*, 1940, **11**, 457–466.
3. KURTZ, E. B., AND LARSEN, M. J., An electrostatic audio generator, *Elec. Engineer.*, 1935 (September).
4. PETERS, W. E., *Die Auffassung der Sprechmelodie*, Leipzig: Theosophischen Verlagshaus, 1924.
5. ROOT, A. R., Pitch-patterns and tonal movement in speech, *Psychol. Monog.*, 1930, **40**, 109–159.
6. SHOWER, E. G., AND BIDDULPH, R., Differential pitch sensitivity of the ear, *J. acoust. Soc. Amer.*, 1931, **3**, 275–287.

(Manuscript received January 12, 1940)

# THE INFLUENCE OF METHODOLOGY UPON THE MEASUREMENT OF REMINISCENCE *

BY SUSAN GRAY

*George Peabody College for Teachers*

The purpose of the investigation reported upon in the present paper has been to evaluate a certain technique customarily used in the measurement of reminiscence and to determine the influence, if any, of this technique upon the type of results obtained.

In studies of reminiscence which employ meaningful verbal material the usual procedure has been to require a reproduction of material immediately after the period of learning and another reproduction after a stated interval. Any increase in retention from a first to a second reproduction of the material has been considered as reminiscence. From the time of Ballard's classic investigation (1) the possibility has been recognized that a written reproduction of material immediately after learning might function as an extension of the learning period. If an immediate recall of material functions in such a manner, improvement in retention obtained with such a technique cannot be considered as reminiscence, especially if reminiscence is defined conventionally as improvement in recall after an interval of time in which there has been no overt practice or rehearsal of material. That the majority of studies of reminiscence with meaningful verbal material have been invalidated by the employment of a technique of double recall has been suggested by Bunch (4) and Deese (5). To the writer's knowledge, however, the influence of the presence or absence of immediate recall upon subsequent recall of meaningful verbal material in studies of reminiscence has not been subjected to any adequately controlled experimentation. This analysis will set forth the results of

* The writer wishes to acknowledge the valuable suggestions and criticisms of Dr. Paul L. Boynton in the preparation of this study.

an experiment which attempted to measure the influence of immediate recall upon delayed recall of material.

## The Experiment

*Subjects.*—The subjects used in the experiment consisted of 88 fifth and sixth grade children, with a median chronological age of 136.3 months, from a public and a private elementary school in Nashville, Tennessee. For each of the three types of verbal material employed in the main experiment the subjects were equated into two groups separately. The basis of division into approximately equal groups was chronological age and raw scores upon an immediate test of retention of the verbal material to be employed. An attempt was made to construct the materials used in matching the groups in such a way as to make them as similar as possible to those used in the main experimental sessions and described in the following section of this analysis.

*Materials.*—Since it is possible that the relative degree of retention may vary with the type of verbal material used, three types of material were employed, ranging from that permitting very little organization to that permitting a relatively high degree. Three forms of each type of material were prepared, all as similar as possible. The first form was used in the preliminary test and the other two in the main experimental sessions.

The first type of material consisted of three lists of 30 words each, chiefly monosyllabic and probably familiar to most fifth and sixth grade children. Ten words illustrative of list B follow.

| | | | | |
|---|---|---|---|---|
| tree | rose | cut | bean | bread |
| rabbit | eye | cherry | think | walnut |

The second type of material was made up of short statements of fact so arranged that although there was a possibility of organization within each sentence, very little organization was possible from sentence to sentence. As an example of the material used, five statements from list B are reproduced as follows:

The lawn mower was invented by a man named Hills.
A cowbird puts its eggs in other birds' nests.
Frederick the Great was a German hero.
There are very few large forests in China.
Turner was a great English painter.

The last type of material consisted of three short, discursive paragraphs, approximately 200 words in length, one each upon the lives of Handel, Constable, and Charles Sturt. An attempt was made to construct each paragraph in such a manner that it would be made up of 21 facts concerning the personage under consideration.

*Method.*—The procedure was that of a group experiment. The method of complete exposure was employed in the learning of the materials, each subject being given a mimeographed copy of the material to be learned. Five minutes was used for the learning period throughout. Ten minutes was allowed for the written reproduction of the words and sentences and twelve minutes for the reproduction of the paragraph. These time allotments for reproduction of the materials proved ample in practically all cases. Verbatim reproduction was required with the words and substance reproduction for the sentences and paragraph. The reproduction of the sentences and paragraphs was scored upon an all or none basis, full credit being given for each factual item correctly reported as to substance and no credit for items incorrectly reproduced or fragmentary.

Coefficients of correlation between immediate recall of the materials employed in the main experimental sessions and a delayed recall one day later ranged from .79 to .93 with probable errors from .04 to .01. These coefficients, based upon 44 cases, were considered large enough to indicate a fairly high and consistent reliability of the method of measuring retention.

Two different conditions of learning and testing were used. The first consisted in a written reproduction immediately after learning and a second reproduction twenty-four hours later. The second consisted in only one reproduction of material given twenty-four hours after learning. During the period allotted for immediate reproduction those subjects who did not give an immediate recall occupied the time in columnar addition of two place numbers presented upon mimeographed sheets.

The first two weeks of the main experiment were employed in the learning and testing of the lists of unrelated words. The first week one group gave both an immediate and a delayed recall of material; the other group gave only a delayed recall. The second week the conditions were reversed. An analogous procedure was followed during the next four weeks with the sentences and the paragraphs. In an attempt to obviate the possibility of the influence of unequal difficulty of material, the different forms of the material used were alternated within each group in such a manner that all forms of material were represented equally in each condition of learning and testing. It may be seen from the procedure outlined that, in general, the method of reversed groups was used; each child was subjected to both conditions of testing employed and gave delayed recalls of each of the three types of material under conditions both of an immediate and of no immediate recall.

## Results

The results of the experiment are derived in the main from the following comparisons:

1. Comparison of median scores of immediate recall of material with median scores of delayed recall which was preceded by immediate recall.
2. Comparison of immediate recall with delayed recall not preceded by immediate recall.
3. Comparison of delayed recall preceded by immediate recall with that not preceded by immediate recall.

Although in the actual handling of results comparisons were made between the two groups when subjected to different conditions of testing and between the scores of each group under one condition with its scores under another condition, only the comparisons between the scores of all cases taken together under one condition and the scores of all cases under another condition will be considered here.

*A. Comparison of Immediate Recall with Delayed Recall Preceded by Immediate.*—It has been with a comparison of

this nature between a primary and a secondary recall of the same material that improvement in retention has been observed with meaningful verbal material and designated as reminiscence. Results of this experiment as indicated in Table 1 show no such improvement when the median of all cases is considered. With words and sentences a highly significant loss from a first to a second recall of material is shown. This is evidenced by critical ratios of 10.5 and 10.7. The loss, however, with the paragraph was so small as to be of very little significance. The results of this comparison, al-

TABLE 1

COMPARISON OF IMMEDIATE RECALL AND DELAYED RECALL PRECEDED BY IMMEDIATE

*A. Group Scores*

| | No. | Immediate Recall | | | Delayed Recall Preceded by Immediate | | |
|---|---|---|---|---|---|---|---|
| | | Mdn | Q | $PE_{mdn}$ | Mdn | Q | $PE_{mdn}$ |
| Words | 88 | 13.50 | 4.72 | .633 | 9.67 | 4.59 | .614 |
| Sentences | 84 | 5.43 | 1.68 | .232 | 3.93 | 1.56 | .214 |
| Paragraph | 88 | 12.38 | 3.27 | .436 | 12.14 | 2.94 | .393 |

*B. Comparison of Differences in Medians*

| | Diff. | $PE_{diff.}$* | C. R. |
|---|---|---|---|
| Words | 3.83 | .366 | 10.46 |
| Sentences | 1.50 | .141 | 10.65 |
| Paragraph | .24 | .167 | 1.43 |

* In this table and in the tables that follow the probable errors of the difference were computed by the formula

$$PE_{diff.} = \sqrt{PE^2_{mdn_1} + PE^2_{mdn_2} - 2r_{12}PE_{mdn_1}PE_{mdn_2}}.$$

though opposed to those of certain investigators employing a double recall technique, as for example, Ballard (**1**), McGeoch (**10, 11**) and Williams (**14**), are in agreement with other investigations which have made use of a similar technique. Examples of such studies are those of Brown (**2**), Gillette (**6**), and Whitely and McGeoch (**13**). The conflicting data obtained with a method of double recall suggest the possibility that certain factors are at work which have not been subjected to control under the experimental procedures employed.

*B. Comparison of Immediate Recall with Delayed Recall not Preceded by Immediate.*—If an immediate recall of material acts as an extension of the learning period, then only in a comparison in which the consideration of two recalls of the same material is eliminated will the presence or absence of reminiscence be demonstrated. Such a comparison is the one given here between immediate recall of one form of material with delayed recall of another form. Not only was there no improvement in recall shown in the group scores, but as indicated in Table 2 the scores were significantly lower through-

TABLE 2

COMPARISON OF IMMEDIATE RECALL WITH DELAYED RECALL NOT PRECEDED BY IMMEDIATE

*A. Group Scores*

| | No. | Immediate Recall | | | Delayed Recall Not Preceded by Immediate | | |
|---|---|---|---|---|---|---|---|
| | | Mdn | Q | $PE_{mdn}$ | Mdn | Q | $PE_{mdn}$ |
| Words | 88 | 13.50 | 4.72 | .633 | 6.43 | 3.79 | .507 |
| Sentences | 84 | 5.43 | 1.68 | .232 | 2.17 | .98 | .136 |
| Paragraph | 88 | 12.38 | 3.27 | .436 | 9.00 | 3.33 | .446 |

*B. Comparison of Differences in Medians*

| | Diff. | $PE_{diff.}$ | C. R. |
|---|---|---|---|
| Words | 7.07 | .699 | 11.09 |
| Sentences | 3.26 | .195 | 16.72 |
| Paragraph | 3.38 | .392 | 8.61 |

out upon the delayed recall. As calculated from median scores, losses in retention were 52.5 percent for words, 60.0 percent for sentences, and 27.4 percent for the paragraph. As with the comparison of immediate recall and delayed recall preceded by immediate, a marked tendency is manifested for the loss in retention to be smaller with material permitting considerable organization than with that permitting little organization.

*C. Comparison of Delayed Recall Preceded by Immediate with that not Preceded by Immediate Recall.*—Probably the most significant results of the experiment in so far as throwing light

upon the influence of immediate recall upon the measurement of reminiscence are those given in Table 3. Here it may be seen that with all three types of material there was a significantly lower median score upon delayed recall not preceded by immediate than upon delayed recall which was preceded by immediate. The only variable in procedure here was that under the one condition a written reproduction was given immediately after learning and under the other condition simple computation was substituted for this immediate written reproduction. The results point clearly to the functioning

TABLE 3

COMPARISON OF DELAYED RECALL PRECEDED BY IMMEDIATE AND DELAYED RECALL NOT PRECEDED BY IMMEDIATE RECALL

*A. Group Scores*

| | No. | Delayed Recall Preceded by Immediate | | | Delayed Recall Not Preceded by Immediate | | |
|---|---|---|---|---|---|---|---|
| | | Mdn. | Q | $PE_{mdn}$ | Mdn. | Q | $PE_{mdn}$ |
| Words | 88 | 9.67 | 4.59 | .614 | 6.43 | 3.79 | .507 |
| Sentences | 84 | 3.93 | 1.56 | .214 | 2.17 | .98 | .136 |
| Paragraph | 88 | 12.14 | 2.94 | .393 | 9.00 | 3.33 | .446 |

*B. Comparison of Differences in Medians*

| | Diff. | $PE_{diff.}$ | C. R. |
|---|---|---|---|
| Words | 3.24 | .654 | 4.97 |
| Sentences | 1.76 | .179 | 9.83 |
| Paragraph | 3.14 | .350 | 8.96 |

of that immediate recall as an extension of the learning period, or, as Warner Brown (3) considers it, as a fixing of the items upon the first test by a written reproduction.

## DISCUSSION

Within limitations of the experimental technique and its errors, the data of this experiment offer some basis for an answer to the major question proposed in this investigation. This question has been: Is the double recall technique, as customarily used in measuring reminiscence, a determiner of such improvement in recall as may be observed?

The findings in the case of all three types of material

indicate that an immediate recall of material does exert a measurable influence upon a second recall of material. These findings, if valid, indicate the questionable nature of studies of reminiscence which have employed a method of double recall. It is not meant by the preceding statement that all rises in retention obtained with a double recall technique may be explained solely on the basis of the functioning of an immediate recall as an extension of the learning period. Rather it is meant that if an immediate recall is demonstrated to function as an additional practice of material, a rise in retention under such a technique may not be considered as reminiscence when reminiscence is defined as improvement after an interval during which there has been no practice or learning of material. A valid technique for the measurement of reminiscence with meaningful verbal material by the complete exposure method would employ some way of obviating the necessity of an immediate test upon the material as a measure of the degree of learning. At least two possible methods may be employed. One method is that of using two carefully equated groups, one of which would give an immediate recall of material and one of which would give a recall after a stated interval of time. Another plan would be to use in double fatigue order four sets of strictly comparable material with the same subjects in such a manner that scores upon an immediate recall of two sets of material might be compared with scores upon delayed tests of the comparable material. A somewhat similar method of employing four sets of comparable material with a group in such a way as to provide for practice equilibration has been employed by Deese (**5**). With such a technique she found no evidence of reminiscence in the total scores of the group tested.

Up to the present time positive evidence of reminiscence with meaningful verbal material has been open to criticism because of the use of the double recall technique. It is the opinion of the writer that the questionable procedures employed and the conflicting results obtained do not permit one to identify improvements in recall obtained by such methods with true reminiscence. Improvement in recall of verbal

material with experimental techniques which controlled rehearsal of material, including an immediate recall, has been obtained by Hovland (**8**, **9**, **10**) and Ward (**12**) with series of nonsense syllables presented by the anticipation method. The rise in retention obtained by Hovland and Ward may be correctly designated as reminiscence in that it is in accordance with the customary definition of the phenomenon. In view of the differences in experimental procedures followed, a possibility arises that improvements in retention obtained with a method of complete exposure and double recall and that improvement obtained by Hovland and Ward are phenomena dependent upon causes entirely different. The results of the present investigation would suggest that the rise in retention obtained in the past with the complete exposure method is not attributable to the presence of true reminiscence but rather is an artifact of the experimental setup.

(Manuscript received January 12, 1940)

## References


1. Ballard, P. B., Obliviscence and reminiscence, *Brit. J. Psychol., Monog. Suppl.*, 1913, **1**, No. 2, pp. vi + 82.
2. Brown, W., Effects of interval on recall, *J. exper. Psychol.*, 1924, **7**, 469–474.
3. ——, To what extent is memory measured by a single recall?, *J. exper. Psychol.*, 1923, **6**, 377–382.
4. Bunch, M. K., The measurement of reminiscence, *Psychol. Rev.*, 1938, **45**, 525–531.
5. Deese, E. L., A study of the retention of meaningful verbal material at short time intervals, *Master's Essay, Yale Univ.*, 1938.
6. Gillette, A. E., Learning and retention: a comparison of three experimental procedures, *Arch. Psychol.*, 1935–1936, **28**, No. 198, p. 56.
7. Hovland, C. I., Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and distributed practice, *J. exper. Psychol.*, 1938, **22**, 201–224.
8. ——, Experimental studies in rote-learning theory. II. Reminiscence with varying speeds of syllable presentation, *J. exper. Psychol.*, 1938, **22**, 338–353.
9. ——, Experimental studies in rote-learning theory. III. Distribution of practice with varying speeds of syllable presentation, *J. exper. Psychol.*, 1938, **23**, 172–190.
10. McGeoch, G. O., The conditions of reminiscence, *Amer. J. Psychol.*, 1935, **47**, 65–89.
11. ——, The factor of degree of learning in reminiscence, a comparative study of preschool children and college students, *J. genet. Psychol.*, 1935, **46**, 455–462.
12. Ward, L. B., Reminiscence and rote-learning, *Psychol. Monog.*, 1937, **49**, No. 4, pp. v + 64.
13. Whitely, P. L. and McGeoch, J. A., The curve of retention for poetry, *J. educ. Psychol.*, 1928, **19**, 471–479.
14. Williams, O., A study of the phenomenon of reminiscence, *J. exper. Psychol.*, 1926, **9**, 368–387.

# AN EXPERIMENTAL INVESTIGATION OF NEGATIVE AFTER-IMAGES OF HALLUCINATED COLORS IN HYPNOSIS *

BY FRANCIS W. HIBLER

*Illinois State Normal University*

The reality of the hallucinatory experience in hypnosis has been questioned frequently. The approach of Binet and Féré (**1**) consisted of suggesting a hallucinated color on a white card and asking the hypnotized subject to fixate a dot on it. After a strong suggestion as to the presence of one of the primary colors the original card was removed leaving the eyes of the subject fixated on a second white card relative to which no suggestion was made. The subject was then asked to report the color seen on the second card. Erickson and Erickson (**2**) adopted the method of presenting 16 plain sheets of white typing paper. It was suggested to the hypotnized subjects that each sheet was one solid bright color. These sheets were then presented one at a time, the experimenter naming the color of the first sheet while the subject named the color of the second sheet, continuing in this alternate fashion. For purposes of control word association tests were used, both before and after the experimental series, to determine any associations the subjects might have made with the four primary colors used in the investigation. Both of these studies report the presence of negative after-images as the result of the suggestion of hallucinated colors in hypnosis.

## Experimental Procedure

I. *Preliminary Control Series.*—The Ishihari test of color discrimination was administered to each subject. Only those individuals who showed normal color vision on the test were used as subjects. Four male subjects were used in the experiment, each of whom met the color vision criterion, and was in addition, a deep trance subject

* This investigation was part of a research which was conducted under the direction of Dr. F. C. Dockeray and Dr. H. H. Goddard in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of The Ohio State University.

showing anesthesia, positive and negative hallucinations, post-hypnotic suggestion and amnesia. All subjects were college students who were without knowledge as to the nature and object of the experiment.

After the Ishihari test had been administered, the extent of the subject's knowledge of negative after-images was checked. The following problem was read to him—

> Suppose that you should stare steadily at a small blue disc, mounted on a white card, for thirty seconds. Now, suppose that I should suddenly remove the blue disc, leaving only the white card; if you should continue to stare at the white card without moving your eyes, would you see anything? What? What color?

The procedure was then repeated, substituting instructions for a red, yellow and a green disc. The experimenter sat out of the field of vision of the subject while the questions were being answered and attempted to give no cue as to their correctness. Frequently, however, the subject was further questioned as to where he obtained his information.

II. *Trance Hallucination of Colors on Grey Cards.*—The subject was hypnotized and permitted to remain in the trance for three minutes.[1] He was then instructed to place his eyes at the eye-piece of the exposure apparatus. The Dockeray tachistoscope was used for this purpose. It permitted the control of intensity of the illumination, peripheral visual field, distance from the subject's eye to the stimulus card, the point of fixation, and the duration of the exposure, as well as providing a convenient method of removing the card with the hallucinated color and leaving the eyes of the subject fixated on a second grey card in the same position. The stimulus color in this series was always hallucinated. Two identical grey cards were placed in the slot at the back of the tachistoscope, one in front of the other. The one exposed to the view of the subject was designated as the stimulus card when the color disc was hallucinated on it. The back one was the screen used for fixation in the production of the after-images. The subject would fixate on the stimulus card for forty-five seconds, after which the experimenter would remove it with a swift motion, exposing the second grey card in the identical position formerly occupied by the stimulus card. This permitted the gaze of the subject to shift from the stimulus card to the second card without moving the eyes, as the fixation points on the two cards were in the same relative position. This technique was found to be efficient in the production of negative after-images with naive subjects.

After the subject was seated, looking into the tachistoscope, the following suggestion was read to him—

> "On that card there is a bright blue disc, ¾ of an inch in diameter. You see it clearly, don't you? All right, now stare steadily at the center of the disc for 45 seconds. (The experimenter repeating instructions of steady fixation.) Now I am going to remove the disc leaving only a grey card. Continue to stare at the dot on it (fixation point) and do not move your eyes. (Experimenter removes stimulus card.) Do you see anything? What? What color?"

The subject was then permitted to rest for three minutes, after which he was again instructed to gaze into the tachistoscope. The blue disc was hallucinated as before, but the following directions were added while the subject was fixating the hallucinated disc.

---

[1] Approximately five minutes were spent in the induction of each trance in this investigation. Only adequately trained subjects were used, and at all times each subject met all the usually accepted criteria of the deep trance. It is interesting to note that Erickson (3) has suggested that a "slow induction extending over 15 to 30 minutes is desirable." The exact importance of this time element is certainly worthy of further research.

"Now I am going to wake you to the count of three.[2] When I say *three* you will be wide awake, feeling fine in every way. While you are awakening your eyes will remain glued on the dot on the card in front of you, and you will not remember anything that happened while you were asleep. Ready, one-two-three. What do you see?" (Experimenter removed card with hallucinated color between count of one and two.)

After giving his report the subject was permitted to rest for five minutes. He was then re-hypnotized and the same procedure was repeated three more times, substituting a hallucinated red, yellow, and green disc for the blue one. After each waking report the extent of the amnesia for the trance event was checked by intensively questioning the subject concerning what happened while he was hypnotized. This concluded the first experimental period.

III. *Trance Hallucinations Involving Color Changes.*—The procedure in this series was identical with that in the foregoing series, with the exception of the nature of the hallucination. The reports were obtained in the same fashion, both in the trance and the waking states.

The nature of the stimulus differed in this respect, that in the former series a color was suggested as being on a blank light grey card, while in this series an actual color disc was always presented, but in each case it was hallucinated as some other color. For example, the first presentation involved the hallucination of a blue disc as a red disc. The method of suggesting the hallucination was as follows. Two tachistoscope cards, one with a red disc, ¾ inch in diameter mounted upon it, and the other with a blue disc of the same size were placed adjacently on the table before the hypnotized subject. The experimenter then said:

"You see this disc, what color is it? (Pointing to the red disc.) That's right. Now this disc (Blue) is exactly the same as this one (Red). It is the same in every way, the same hue, brightness, and saturation. It is the *same* color. Now take this card in your hand and describe this disc to me. (Subject takes blue disc and describes it as red, giving size and shade as a rule.) That's right, now remember, this (Blue) is the same as this (Red) until I tell you differently."

The subject was then instructed to look into the eyepiece of the tachistoscope and the blue disc was presented. The subject was asked to describe what he saw, and on describing the blue disc as red, was given the same instructions as in the former series. After the final waking report, the subject was re-hypnotized and the suggestion was removed, after which he was checked for normal color vision. Ten minutes rest was given between the presentation of one color and the next.

The procedure of this series was repeated four times, once with each of four suggestions of color changes. These hallucinations were, (1) blue as red; (2) green as yellow; (3) red as blue; (4) yellow as green. Only the four primary colors were used, and in no case was a color hallucinated as its complementary. This was to guard against a possible positive after-image of the real color being reported as the negative after-image of the hallucinated color. For example, if green was hallucinated as red, a reported after-image of green might be either the positive after-image of the genuine green, or it could be the negative after-image of the hallucinated red. At no time, however, did a positive after-image appear to complicate the results. It required an entire experimental period to complete this series.

---

[2] A preliminary training series was used to make sure all subjects would awaken at the count of three.

IV. *Final Control Series (Trance and waking state).*—The third experimental period for each subject was used in completing the final control series. The first step was a further check of the subject's knowledge as to complementary colors. This was accomplished by asking the same questions as in the preliminary control, "What would you see if you gazed steadily at a blue (red, yellow, green) disc, and it was suddenly removed, leaving you looking at a white card?" This provided an additional check on the extent of the post-hypnotic amnesia, inasmuch as the only way the subject's original idea of after-images could have been changed, was by remembering the trance color and either the trance or waking report of the after-image produced by it.

As still an additional check on whether or not the subject had been able to remember the trance presentation of the real colors, and derive any principles as to the laws of complementary colors, the entire procedure described under 'trance hallucination of colors on grey cards' was repeated. It is reasonable to assume that if the subject did not know the color stimulus that produced the images experienced in the waking state, he would have no reason to react differently on the second presentation of the series than on the first. If, however, the hallucinations involving changes in real colors had increased his sophistication with respect to after-images, this would possibly be disclosed in his reactions to the series when it was presented the second time. In any event, any change in responses between the two presentations of the same series would be significant to the point of requiring further scrutiny of the technique used.

The final step of the control series was the presentation of the actual color discs in the trance state without suggestion, and in the waking state. Each of the four colors used in the investigation—red, green, blue and yellow—were presented in each state, and the report of the after-image experienced from each was noted. This was deemed necessary in order to validate the apparatus used, but was placed last in the experiment so as not to influence the reports with respect to the hallucinations.

## Results of the Experiment

The results of the experiment are summarized in Tables I and II. In addition to the tabular results of the experiment, the protocols taken on the subjects with respect to the experienced hallucinations are significant in some instances.

*Subject A.*—This subject stated that after-images would appear the same as the color presented. To each color hallucinated on a blank grey card, he reported as an after-image in the trance just what the experimenter suggested as the hallucinated color. For him, trance after-images were relatively permanent and identical in every way with the hallucinated stimulus color. The waking report, however, was characteristically—'Just a white square with a pin hole in the center of it.'

In the trance hallucinations involving color changes he again gave his report of the trance after-image as identical with the hallucinated color. His waking reports of the after-

## TABLE I

### Control Series

| Subject | Color Stimulus | Subject's Expectation of After-Image | |
|---|---|---|---|
| | | Preliminary Series | Final Re-check Series |
| A | Blue | Blue | Blue |
| B | | Orange | Orange |
| C | | Yellow | Yellow |
| D | | Yellow | Yellow |
| A | Red | Red | Red |
| B | | Green | Green |
| C | | Green | Green |
| D | | Green | Green |
| A | Yellow | Yellow | Yellow |
| B | | Purple | Purple |
| C | | Blue | Blue |
| D | | Blue | Blue |
| A | Green | Green | Green |
| B | | Red | Red |
| C | | Red | Red |
| D | | Red | Red |

### Final Control Series

| Subject | Color Stimulus (Hallucinated or actual) | Presentation of *Hallucinated* Color Stimulus | | Presentation of *Actual* Color Stimulus | |
|---|---|---|---|---|---|
| | | Trance After-image | Waking After-image | Trance After-image | Waking After-image |
| A | Blue | Blue | None | Yellow | Pale Yellow |
| B | | Orange | None | Orange | Orange |
| C | | Yellow | None | Yellow | Yellow |
| D | | Yellow | None | Yellow | Yellow |
| A | Red | Red | None | Green | Green |
| B | | Green | None | Light Green | Light Green |
| C | | Green | None | Light Green | Light Green |
| D | | Green | None | Pale Green | Green |
| A | Yellow | Yellow | None | Blue | Purple |
| B | | Purple | None | Purple | Purple |
| C | | Blue-grey | None | Blue | Bluish Purple |
| D | | Blue | None | Blue | Blue |
| A | Green | Green | None | Pink | Pink |
| B | | Red | None | Red | Light Red |
| C | | Red | None | Pink | Pink |
| D | | Red | None | Pink | Pink |

TABLE II

EXPERIMENTAL RESULTS

| Subject | Trance Hallucination of Colors on Blank Cards | | | Trance Hallucination Involving Color Changes | | |
|---|---|---|---|---|---|---|
| | Color Hallucinated as Stimulus | After-image Reported in Trance | After-image Reported in Waking State | Suggested Hallucination | After-image Reported in Trance | After-image Reported in Waking State |
| A | BLUE | Blue | None | RED AS BLUE | Blue | Green |
| B | | Orange | None | | Green | Green |
| C | | Yellow | None | | Yellow | Green |
| D | | Yellow | None | | Green | Green |
| A | RED | Red | None | BLUE AS RED | Red | Orange |
| B | | Green | None | | Orange | Orange |
| C | | Green | None | | Green | Yellow |
| D | | Green | None | | Yellow | Yellow |
| A | YELLOW | Yellow | None | GREEN AS YELLOW | Yellow | Pink |
| B | | Purple | None | | Red | Red |
| C | | Blue-grey | None | | Blue | Pink |
| D | | Blue | None | | Pink | Pink |
| A | GREEN | Green | None | YELLOW AS GREEN | Green | Purple |
| B | | Red | None | | Blue | Blue |
| C | | Red | None | | Red | Purple |
| D | | Dark red | None | | Blue | Blue |

images, however, were quite different. For example, his waking report of the hallucination of blue as red was—'I saw an orange color in there; it was sort of hazy but was in the general shape of a disc.' His general behavior indicated that he was perplexed at what he saw, and on one occasion he asked, 'What makes that color disappear that way?' At the end of the experiment he stated that at no time did he realize that the colors he saw were after-images.

In the final control series when the actual colors were presented without suggestion, all of his responses were definite. For example, his trance report of the after-image of the blue disc was, 'I see a yellow disc,' while his waking report was, 'Now I see a bright yellow disc.' An analysis of the responses in this control series does not reveal any consistent or significant differences between trance and waking report of after-images.

*Subject B.*—This subject understood the laws of negative after-images. Hence his report of the after-image of a blue

disc hallucinated on a plain grey card is significant when compared to that of Subject A. Subject B reported it as follows: "It's a faint orange; it's moving a little; it's almost a yellow; now it's gone." A's report was highly *artificial* while the report of B was an excellent description of a negative after-image. Each of his descriptions of after-images to hallucinated colors was most realistic. In the waking report of the after-image to the trance hallucination, however, the characteristic response was, 'Just a white card with a pin point on it.' On one occasion he remarked, "Nothing but the white card with the dot. Are you fooling me today or something?"

In the hallucinations involving color changes this subject always reported the after-image of the real, instead of the hallucinated, color. This held true for both the trance and waking reports, despite the fact that the hallucination appeared to be entirely functional in the trance. Consequently there is no significant difference between trance and waking reports.[3]

*Subject C.*—Subject C also knew the laws of complementary colors, but he stated in the preliminary control series that he was not at all sure whether one would see positive or negative after-images in the set-up described. Considering this set, his trance report of the after-image to blue hallucinated on a plain card is significant, "A hazy blue disc changing to yellow; getting more yellowish and slightly reddish around the outside. Now it's yellow and beginning to fade." It will be noticed that this subject brought both the positive and negative after-image into his report. He did this in each instance of a hallucination of a color on a plain grey card, but in no case where there was an actual color presented. Under the latter circumstances his report was always clear-cut, even when he was describing the trance after-image of hallucinations involving color changes. For example, with blue hallucinated as red, his trance report was, "A bright green hazy

[3] In explaining this experiment to the subject after it was completed, he remarked that he did not think the color discs he saw on awakening were after-images because they did not appear to move. This suggests the possibility that post-hypnotic suggestion may be a valuable technique in the prevention of eye movements as it appeared effective in this instance.

disc; shape is changing, getting irregular and larger." His waking after-image of the same trance stimulus was, 'Yellow after-image; rather irregular, round, about an inch in diameter.' This was the only subject to recognize the color seen in the tachistoscope as an after-image before the final control series. He did not, however, succeed in determining how the stimulus was presented, and stated that he was not sure that he saw an after-image at all, but merely believed he did because it disappeared. In the waking control series there were no marked differences between trance and waking reports. As shown in Table II, this subject always reported the proper negative after-image for each hallucinated color, when reporting in the trance state, and each of his reports was vivid and detailed. In the waking report of the trance hallucinations he always reported the correct after-image of the actual color and with language of equal realism.

*Subject D.*—This subject also was sophisticated with respect to complementary colors. His reports rivaled Subject C's for realism and apparent accuracy. For example, his trance report of the after-image of a red disc hallucinated on a blank grey card was as follows: "I see a pale green disc; it's right in the center; it's blurred at the edges; now it's moving." The waking report of the same trance stimulus was, 'Just a card with a little hole in it again.' In general the reports of this subject were also characterized by a marked realism.

## Summary of Results

1. Trance after-images were reported for each case of trance hallucination of colors on plain grey cards.

2. In each instance the after-image so reported was identical with the preconceived ideas of the subject with respect to after-images.

3. When the report of a similarly presented hallucination was made in the *waking* state, no after-image of any sort was ever reported.

4. After-images of trance hallucinations involving color changes were always reported in both the trance and waking states.

5. The after-images in the *trance* reports varied—subjects A and C always reporting the complement of the hallucinated color, while subjects B and D always reported the complement of the actual color.

6. The after-images in the *waking* reports were always the complement of the actual color for all subjects.

7. No subject changed his preconceived ideas with respect to after-images during the course of the experiment.

8. A repetition of the series involving the hallucination of colors hallucinated on plain cards resulted in reports identical with those obtained from the first presentation.

9. All subjects reported entirely normal after-images in both the waking and trance states when actual colors were presented in the tachistoscope without any suggestion of hallucination.

## Discussion

The fact that trance after-images were reported in each case of hypnotic hallucination of a color on a blank grey card, coincides with the findings of previous investigators. In one respect, however, our findings do not agree. In every case others (**1**, **2**) reported that their subjects experienced negative after-images. In our results three subjects reported negative after-images, while one reported an image identical to the suggested hallucination.

The possibility arises that the latter, subject A, experienced a positive after-image. This conclusion, however, seems most improbable in the light of the total data obtained from this subject. In the first place, his description of the after-image was not accurate with respect to the nature of positive after-images in general—they were too clear, too permanent, and too stationary. For example, his after-image of an hallucinated blue disc was reported as follows: "I see a blue disc the same size, and the same color, and in the same place as the other one." No after-image in the final control series, where the actual color was presented without suggestion of hallucination, was so reported. Even more significant than the nature of his descriptions, were his reports of the after-

images in the final control series. In each case he reported a *negative* after-image of the actual color presented; never the positive. Hence we reject, as the explanation of Subject A's trance report of after-images of trance hallucinated colors, the possibility that he experienced a positive after-image.

The question of the reality of the negative after-images experienced by subjects B,C, and D, is important. Their descriptions of the after-images reported are entirely consistent with our knowledge of the nature of negative after-images; they were vague, blurred towards the edges; they moved about; they faded. If the experiment ended with their reports, the investigator must be convinced of the reality of the reported phenomenon. Previous investigators (1) considered their experiment complete at this point and drew their conclusion that the hallucinated object was perceived as if it were a real object. The writer can well understand their conclusions in the light of his data, if the controls were eliminated. However, when a careful control series is used, the reality of these negative after-images must be considered further.

The fact that no after-image of any sort was reported when the hallucinated color was presented in the trance and the report as to the after-image was made in the waking state, should be interpreted. One possible explanation is that the set of waking is enough to destroy the trance after-image. This hypothesis must be discarded, however, in the face of the evidence presented by the later control series where the after-image of an actual color was always correctly reported in the waking state when the same technique of trance presentation, and waking report, was employed. This leaves as the only possible interpretation that if an after-image of any sort is experienced in the trance state as the result of an hallucinated color, it must be very different from that experienced from the presentation of an actual color stimulus. It is different in that—first, it varies with the knowledge of the subject as to the laws of complementary colors; second, it does not persist over into the waking state. In addition, we must

consider the additional data obtained by the use of hallucinations involving color changes.

It will be noted in Table II that two subjects reported that the trance after-image was the complement of the hallucinated color, while the other two subjects reported the after-image in trance was the complement of the actual color. Identical suggestions and color stimuli were given to all subjects, yet the after-images varied, according to reports of the different subjects. That this variation was not due to the experimental set-up itself, is indicated by the results of the final control series, where the same color discs were presented in trance and waking, but without suggestion of hallucination, and the same after-images were consistently reported by all subjects. In the hallucinations involving color changes a more complex situation is presented to the subject (in contrast to hallucinating a color on a blank card) only if the hallucination is not actually seen as real. In other words, if the subject actually experiences red, it should make no difference if that red is super-imposed over blue, or is on a grey card. If, however, the hallucinated red is not actually experienced as red, then a blue background could conceivably be more confusing than one of plain grey. The data show that when the hallucination was presented on a plain grey background, each subject consistently reported the trance after-image to be identical with his preconceived idea of after-images; but when the hallucination was presented over another color, the trance reports of after-images varied—half of the subjects seeing the complement of the hallucinated color, while the other half entirely ignored the hallucinated color to report the complement of the actual color. In the simple hallucination on a plain background, all subjects were consistent in their reports; in the more complex hallucination of one color over the other, there was a confusion of reports, all of which indicates that the experience of hallucinated 'red' was not consistent in the two types of hallucinations. We interpret this as one indication that the experienced hallucination is not always, if ever, seen as real.

Another question rising out of the trance hallucinations involving color changes is: "If the subject actually sees a blue as a red, why should he *ever* get the after-image of blue, even in the waking state?" This question does not seem possible to answer, when one considers the nature of after-images, and yet all subjects always report the complement of blue under the above conditions. One is thus forced to the conclusion that the important color with reference to after-images is the actual, not the hallucinated color.

In summary, the study of negative after-images has added the following pertinent facts to our general knowledge of positive visual hallucinations:

1. A pre-trance knowledge of complementary colors does influence the report of the after-image of trance hallucinated colors.

2. In trance, with knowledge of the hallucinated color stimulus, a definite after-image is *always* reported; in the waking state, without knowledge of the hallucinated color stimulus, no after-image is ever reported.

3. If one color is trance hallucinated as another, the after-image reported in the trance may be the complement of either the hallucinated or the actual color; the after-image reported in the waking state is always the complement of the actual color.

4. Within the limits of the experiment, there is no evidence that hallucinations produce any central or sensory change in the organism that cannot be explained in terms of mere verbal agreement, and coöperation, with the experimenter.

Our experimental findings lead us to submit the following theoretical position: since hallucinated colors definitely do not have the properties of actual colors, and since no evidence of the existence of the hallucination was discovered except a mere verbal agreement with the hypnotizer, hypnotic hallucinations are probably verbal in nature rather than sensory or a function of the central nervous system. This is the most parsimonious theoretical position regarding the nature of positive visual hallucinations.

Further investigations on this problem are now in the process of preparation.

(Manuscript received January 3, 1940)

REFERENCES

1. BINET, A., AND FÉRÉ, C., *Animal Magnetism* (Kegan Paul, Trench, Trubner & Co., Ltd., London. 1891).
2. ERICKSON, M. A., AND ERICKSON, E. M., The hypnotic induction of hallucinatory color vision followed by pseudo negative after-images, *J. exper. Psychol.*, 1938, **22**, 581–588.
3. ERICKSON, M. H., The induction of color blindness by a technique of hypnotic suggestion, *J. gen. Psychol.*, 1939, **20**, 61–89.

# THE EFFECT OF STIMULATION BY ODOROUS SUBSTANCES UPON THE AMOUNT OF SECRETION OF THE PAROTID GLANDS *

BY CHARLES A. ELSBERG, M.D., H. SPOTNITZ, M.D., AND E. I. STRONGIN, PH.D.

In the investigations here reported, the effects of the stream injection of odors upon the secretory activity of the parotid glands were studied.

The technique used for the stimulation with odors was that described in a paper recently published (1). The apparatus and technique used for the collection and measurement of the parotid secretion were essentially those of Winsor and his associates (2), and of Strongin and his co-workers (3).

It is sufficient to indicate that a small disk similar to the one developed by Lashley is fastened by suction over the opening of each Stenson's duct. From the inner chamber of each disk, a small air-tight tube passes through the corner of the mouth and carries the parotid secretion to a calibrated glass tube marked off in hundredths of a cubic centimeter. We have sometimes used calibrated tubes marked off in thousandths of a cubic centimeter.

The subject was seated on an easy chair and instructed to assume a comfortable position. In order to avoid distracting influences, the tests were made in a partly sound-proof room in which there were only the examiner and his subject.

After the disks had been applied to the openings of Stenson's ducts and the conducting tubes had been connected to the calibrated glass tubes, a period of ten minutes was allowed to elapse before the stream injection with the odorous substance was begun. During this ten-minute period, the amount of secretion from each parotid gland was recorded every thirty seconds. The object of this preliminary period was to allow the individual to become adapted to the presence of the disks and to exclude, as far as possible, any effect upon the secretory activity of the parotid glands due to the application of the disks.

The individual then received a stream injection of odor (4) into one or into both nasal passages for a period of five minutes. During this time, the amount of secretion from each parotid gland was recorded at thirty-second intervals. After the stimulation was concluded, the amount of secretion during the next five minutes was also recorded.

The figures given in Table I, Column 2, are the average amounts secreted *each minute* by the two parotid glands in 50 normal individuals without stimulation, and in Column 3, the *total amounts* of the secretion from the two parotid glands without stimulation.

---

* From the Clinical Research Laboratories of the Neurological Institute of New York.

Work done under a grant for brain tumor research from the John and Mary R. Markle Foundation.

In Tables II and III are given the amounts of parotid secretion from each gland and from the two glands together during a stream injection of air (Table II) and of the odor of coffee (Table III).

In Table IV, Column 3, are given the total amounts of secretion from the *left* parotid gland during the stream injection of odor into the *left* nasal passage.

The figures in Tables V, VI and VII (in Column 2 in each table) are the total amounts secreted by the parotid glands of the two sides during the bilateral stream injection of odor.

In the normal individual, the rate of secretion of the parotid glands is a regular one, as shown by Winsor, Strongin and Korchin.[1] We have used their records of the amounts of parotid secretion of fifty normals recorded every minute for the calculation of the average secretory rate.

Table I shows that for the first three to four minutes, the rate of secretion was disturbed by some factor (possibly the gradual adaptation to the presence of the disks). Thereafter, the rate of parotid secretion remained fairly constant (Table I, Column 4).

TABLE I

AVERAGE SECRETION OF THE TWO PAROTID GLANDS (50 INDIVIDUALS)

| | Average per Individual | | |
|---|---|---|---|
| | Amount of Secretion from Both Parotid Glands Each Minute $\left(\frac{\text{cc.}}{100}\right)$ | Total Amount $\left(\frac{\text{cc.}}{100}\right)$ | Total Amount of Secretion ÷ Time |
| 1st minute | 5.18 | 5.18 | 5.18 |
| 2nd minute | 4.42 | 9.60 | 4.80 |
| 3rd minute | 4.66 | 14.26 | 4.74 |
| 4th minute | 3.78 | 18.04 | 4.51 |
| 5th minute | 3.46 | 21.50 | 4.30 |
| 6th minute | 4.70 | 26.20 | 4.33 |
| 7th minute | 3.82 | 30.02 | 4.51 |
| 8th minute | 4.24 | 34.26 | 4.28 |
| 9th minute | 4.46 | 38.72 | 4.30 |
| 10th minute | 4.40 | 43.10 | 4.31 |
| 11th minute | 5.26 | 48.36 | 4.38 |
| 12th minute | 3.64 | 52.00 | 4.33 |
| 13th minute | 4.02 | 56.02 | 4.31 |
| 14th minute | 4.30 | 60.32 | 4.31 |
| 15th minute | 2.84* | 63.16 | 4.21 |

* During the last five-minute period of these tests, the average total amount per individual, secreted by the two parotid glands was 20 hundredths of a cc.

[1] We are indebted to Drs. Winsor and Korchin for a large number of records of tests on normal individuals.

The effect of a stream injection of *air* at a moderate pressure and rate (1000 cc. to 4000 cc. per minute) does not distinctly increase the amount of secretion of the parotid glands (see Table II).

TABLE II

C.A.E. STREAM INJECTION OF AIR AT THE RATE OF 2000 CC. PER MINUTE INTO RIGHT NASAL PASSAGE FOR 5 MINUTES

| Duration of Stream Injection of Air into Right Nasal Passage (seconds) | Total Secretion from Parotid Glands ($\frac{cc.}{100}$) | | Total Secretion from Both Parotid Glands ($\frac{cc.}{100}$) |
|---|---|---|---|
| | Right | Left | |
| 30 | 0 | 0 | 0 |
| 60 | 2 | 4 | 6 |
| 90 | 3 | 9 | 12 |
| 120 | 5 | 10 | 15 |
| 150 | 6 | 12 | 18 |
| 180 | 7 | 13 | 20 |
| 210 | 7 | 14 | 21 |
| 240 | 8 | 15 | 23 |
| 270 | 9 | 16 | 25 |
| 300 | 10 | 16 | 26* |

* Total amount secreted by the two parotid glands in five minutes, 26 hundredths of a cc. During the preceding five-minute period (without stream injection of air) the total amount secreted was 28 hundredths of a cc. (13 on right, 15 on left).

The effect of a stream injection of air and *odor* varies from individual to individual, and with the odor used for the tests. We have made series of tests with many odors and have injected the stream of air and odor at different stream rates and pressures. Apparently, odors which have little if any effect upon the trigeminal nerve (**5**) such as coffee or phenyl ethyl alcohol for example, have an uncertain effect. Frequently they do not increase the amount of secretion of the parotid glands, or, if there is any effect, it is slight and inconsistent (see Table III).

On the other hand, odorous substances which have a distinct trigeminal effect, such as citral, menthol and turpentine, cause a marked increase of parotid secretion (Tables IV, V, VI and VII). When odors which also effect the trigeminus are used for the tests, *the effect of a stream injection of air and odor into one nasal passage, is a marked increase of parotid secretion from the ipsilateral parotid gland. The effect of a*

TABLE III

G.V. STREAM INJECTION OF AIR SATURATED WITH THE ODOR OF COFFEE AT THE RATE OF 2000 CC. PER MINUTE INTO THE LEFT NASAL PASSAGE FOR 5 MINUTES

| Duration of Stream Injection of Coffee into Left Nasal Passage (seconds) | Total Secretion from Parotid Glands ($\frac{cc.}{100}$) | | Total Secretion from Both Parotid Glands ($\frac{cc.}{100}$) |
|---|---|---|---|
| | Right | Left | |
| 30 | 3 | 1 | 4 |
| 60 | 4 | 2 | 6 |
| 90 | 4 | 2 | 6 |
| 120 | 5 | 2 | 7 |
| 150 | 6 | 4 | 10 |
| 180 | 7 | 5 | 12 |
| 210 | 8 | 6 | 14 |
| 240 | 9 | 7 | 16 |
| 270 | 10 | 7 | 17 |
| 300 | 12 | 8 | 20* |

* Total amount secreted by the two parotid glands in five minutes, 20 hundredths of a cc. During the five-minute period preceding the stream injection, the total amount secreted was 9 hundredths of a cc.

*bilateral stream injection is a marked increase of secretion of both parotid glands.*

Protocols of four series of tests are given in Tables IV to VII.

TABLE IV

G.V. STREAM INJECTION OF AIR SATURATED WITH THE ODOR OF CITRAL AT THE RATE OF 1000 CC. PER MINUTE INTO THE LEFT NASAL PASSAGE

| Duration of Stream Injection of Citral into Left Nasal Passage (seconds) | Total Secretion of Parotid Glands ($\frac{cc.}{100}$) | | Stimulation Time$^{2/3}$ | Amount of Secretion ÷ Stimulation Time$^{2/3}$ | Amount of Secretion Calculated on Basis of $C = 0.398$ |
|---|---|---|---|---|---|
| | Right | Left | | | |
| 30 | 0 | 3 | 9.655 | 0.311 | 3.8 |
| 60 | 0 | 5 | 15.33 | 0.326 | 6.1 |
| 90 | 0 | 8 | 20.08 | 0.398 | 8.0 |
| 120 | 0 | 10 | 24.33 | 0.411 | 9.7 |
| 150 | 0 | 11 | 28.23 | 0.389 | 11.2 |
| 180 | 0 | 12 | 31.88 | 0.376 | 12.7 |
| 210 | 1 | 13 | 35.33 | 0.368 | 14.1 |
| 240 | 1 | 14 | 38.62 | 0.363 | 15.4 |
| 270 | 1 | 16 | 41.70 | 0.384 | 16.6 |
| 300 | 2 | 18* | 44.82 | 0.402 | 17.8 |

* During the preceding five-minute period, without odor stimulation, the total amount secreted by the left parotid gland was 4 hundredths of a cc.

TABLE V

C.A.E. Bilateral Stream Injection of Air Saturated with the Odor of Citral at the Rate of 2000 cc. per Minute

| Duration of Stream Injection of Citral into Both Nasal Passages (seconds) | Sum Total Secretion of Both Parotid Glands $\left(\frac{cc.}{100}\right)$ | Stimulation Time$^{2/3}$ | Amount of Secretion ÷ Stimulation Time$^{2/3}$ | Amount of Secretion Calculated on Basis of $C = 0.647$ |
|---|---|---|---|---|
| 30 | 3 | 9.655 | 0.311 | — |
| 60 | 9 | 15.33 | 0.587 | 9.9 |
| 90 | 13 | 20.08 | 0.647 | 13.0 |
| 120 | 17 | 24.33 | 0.649 | 15.7 |
| 150 | 19 | 28.23 | 0.673 | 18.3 |
| 180 | 22 | 31.88 | 0.690 | 20.6 |
| 210 | 23 | 35.33 | 0.651 | 22.9 |
| 240 | 25 | 38.62 | 0.647 | 25.0 |
| 270 | 28 | 41.70 | 0.671 | 27.0 |
| 300 | 30* | 44.82 | 0.669 | 29.0 |

* During the preceding five-minute period, without odor stimulation, the total amount secreted by the two parotid glands was 19 hundredths of a cc.

TABLE VI

E.I.S. Bilateral Stream Injection of Air Saturated with the Odor of Citral at the Rate of 2000 cc. per Minute

| Duration of Stream Injection of Citral into Both Nasal Passages (seconds) | Sum Total Secretion of Both Parotid Glands $\left(\frac{cc.}{100}\right)$ | Stimulation Time$^{2/3}$ | Amount of Secretion ÷ Stimulation Time$^{2/3}$ | Amount of Secretion Calculated on Basis of $C = 1.17$ |
|---|---|---|---|---|
| 30 | 11 | 9.655 | 1.14 | 11.3 |
| 60 | 18 | 15.33 | 1.17 | 17.9 |
| 90 | 23 | 20.08 | 1.15 | 23.5 |
| 120 | 29 | 24.33 | 1.19 | 28.5 |
| 150 | 34 | 28.23 | 1.20 | 33.0 |
| 180 | 38 | 31.88 | 1.19 | 37.3 |
| 210 | 41 | 35.33 | 1.16 | 41.3 |
| 240 | 44 | 38.62 | 1.14 | 45.2 |
| 270 | 48 | 41.70 | 1.15 | 48.8 |
| 300 | 52* | 44.82 | 1.16 | 52.4 |

* During the preceding five-minute period, without odor stimulation, the total amount secreted by the two parotid glands was 21 hundredths of a cc.

These tables show that—probably because of an additional factor due to the first effect of the injection of air and odor—the formula reported did not give a constant until 30 seconds or one minute had elapsed.

TABLE VII

C.A.E. BILATERAL STREAM INJECTION OF AIR SATURATED WITH THE ODOR OF MENTHOL AT THE RATE OF 2000 CC. PER MINUTE

| Duration of Stream Injection of Menthol into Both Nasal Passages (seconds) | Sum Total Secretion of Both Parotid Glands $\left(\frac{cc.}{100}\right)$ | Stimulation Time$^{2/3}$ | Amount of Secretion ÷ Stimulation Time$^{2/3}$ | Amount of Secretion Calculated on Basis of $C = 1.04$ |
|---|---|---|---|---|
| 30 | 8 | 9.655 | 0.829 | — |
| 60 | 16 | 15.33 | 1.04 | 15.9 |
| 90 | 23 | 20.08 | 1.15 | 20.9 |
| 120 | 28 | 24.33 | 1.15 | 25.3 |
| 150 | 31 | 28.23 | 1.09 | 29.4 |
| 180 | 34 | 31.88 | 1.07 | 33.2 |
| 210 | 37 | 35.33 | 1.05 | 36.7 |
| 240 | 39 | 38.62 | 1.01 | 40.2 |
| 270 | 42 | 41.70 | 1.01 | 43.4 |
| 300 | 47* | 44.82 | 1.05 | 46.6 |

* During the preceding five-minute period, without odor stimulation, the total amount secreted by the two parotid glands was 18 hundredths of a cc.

## COMMENT

It is significant that the average normal secretory rate of the parotid glands is constant in physical time just as, for example, the respiratory rate and the cardiac rate. Such a relationship has the exponent of one ($n = 1$) in the simple power formula

$$(Time)^n \div Total\ Parotid\ Secretion = Constant. \quad (1)$$

In the experiments here described, when a stream of an odor such as citral or menthol was injected into the nasal passages at relatively moderate pressure, the normal secretory rate of the parotid glands was increased. The increase in the rate of secretion is indicated in the power formula by the change of the exponent $n$ from $n = 1$ to $n = 2/3$, *i.e.*

$$(Stimulation\ Time)^{2/3} \div Total\ Volume\ of\ Parotid\ Secretion = Constant. \quad (2)$$

The above equation adequately describes the altered conditions and enabled one to predict the total amount of secretion during the five-minute period of olfactory stimulation from one reading during the experiment.[2]

[2] Experiments, in which the duration of stimulation of the olfactory membranes was longer than five minutes have not yet been made.

Formula 2 may be of importance for an understanding of both the physiology and the psychology of the nervous system.[3]

As has already been reported for three special senses—smell, vision and hearing—the same type of formula with the same exponent

$$(Stimulation\ Time)^{2/3} \div Recovery\ Time = Constant$$

applied to the relationship between the time an individual was stimulated with a relatively moderate stimulus and the time that elapsed after cessation of the stimulus before a threshold stimulus could again be recognized. This so-called recovery time for the special senses is in a way equivalent to the total volume of parotid secretion as a result of stimulation. Both recovery time, and total parotid secretion in the experiments here described indicate the total effects produced by stimulation of the specialized peripheral receptors—in the one instance measured by the time required for return of the ability to recognize a threshold stimulus, in the other measured by the effect on the parotid glands as indicated by the total amount of secretion.

It is significant that essentially the same formula is applicable on the one hand to the effects of stimulation of a peripheral receptor as based upon the time required for the return of a conscious voluntary process (the recognition and reaction to a threshold stimulus), on the other hand to the objective effects of stimulation of a peripheral receptor upon an unconscious involuntary process (parotid secretion).

## Summary

1. The average normal secretory rate of the parotid glands of 50 normal individuals was fairly constant.

2. The volume of fluid secreted by *one* parotid gland is increased by a stream injection of citral or menthol into the *ipsilateral* nasal passage.

[3] Crozier and Holway (**6**) have recently confirmed the fact that the power formula holds for certain visual mechanisms. They express the formula in logarithmic form as

$$\log \Delta I = - Z \log A + C$$

They state that "the properties of $\Delta I$ are quantitatively determined by events central to the retina."

3. The volume of fluid secreted by the *two* parotid glands is increased by stream injection of odorous substances such as citral or menthol into the *two* nasal passages.

4. The total volume of parotid secretion during a five-minute period of a stream injection of odorous substances such as citral or menthol into the two nasal passages was increased according to the formula

$$Stimulation\ Time^{2/3} \div Total\ Amount\ of\ Parotid\ Secretion = Constant.$$

(Manuscript received December 21, 1939)

## References

1. Elsberg, C. A. and Spotnitz, H., Relation of stimulation time of receptors to recovery time in the nervous system: visual, olfactory and auditory senses, *J. Neurophysiol.*, 1939, **2**, 227–233.
2. Winsor, A. L., Conditions affecting human parotid secretion, *J. exper. Psychol.*, 1928, **11**, 355–363.
3. Strongin, E. I. and Hinsie, L. E., Parotid secretory rate in schizophrenic patients, *J. Nerv. and Ment. Dis.*, 1938, **87**, 715–724.
4. Elsberg, C. A. and Brewer, E. D., The sense of smell. X. A detailed description of the technique of two olfactory tests used for the localization of supratentorial tumors of the brain, *Bull. Neurol. Inst. N. Y.*, 1935, **4**, 501–510.
5. Elsberg, C. A., Levy, I. and Brewer, E. D., The sense of smell. VI. The trigeminal effects of odorous substances, *Bull. Neurol. Inst. N. Y.*, 1935, **4**, 270–285.
6. Crozier, W. J. and Holway, A. H., Theory and measurement of visual mechanisms. III. $\Delta I$ as a function of area, intensity and wave length, for monocular and binocular stimulation, *J. gen. Physiol.*, 1939, **23**, 101–140.

# COMPARATIVE RETENTION OF OPEN AND CLOSED VISUAL FORMS

BY H. GURNEE, B. E. WITZEMAN, AND M. HELLER

It is a principle of Gestalt psychology that incompletion affects retention. The effect is said to be a function, not of incompletion as such, but of the amount of tension induced in the system by the circumstances of incompletion. Thus, according to Koffka,[1] where no tension toward completion is induced, memory traces of open systems are likely to be less stable and have lower survival value than memory traces of closed systems of a comparable sort; but where such tension toward completion is induced, the open system will manifest greater survival value than the closed system. Appearing to support this latter principle are several experiments by associates[2] of Lewin, demonstrating a greater recall value for interrupted than for completed tasks. However, recall was here indicated either by naming the task or by resumption of it, and unfortunately these two procedures give no measure of the comparative availability of the systems of behavior whereby the tasks were being performed. For all we know, the response-pattern may have been just as well retained in the completed as in the interrupted task. McKinney[3] seems to have been the only person to investigate this problem in a crucial way, and he reported no clear advantage in favor of the interrupted system.

McKinney used mazes. Visual forms offer another means of testing these Gestalt generalizations. Such forms lend

[1] KOFFKA, K., *Principles of Gestalt Psychology*, 1935, pp. 339–340, 508, 621–622.

[2] ZEIGARNIK, B., Über das Behalten von erledigten und unerledigten Handlungen, *Psychol. Forsch.*, 1927, **9**, 1–85. M. OVSIANKINA, Die Wiederaufnahme unterbrochener Handlungen, *Psychol. Forsch.*, 1928, **11**, 302–379. D. L. ADLER and J. S. KOUNIN, Some factors operating at the moment of resumption of interrupted tasks, *J. Psychol.*, 1939, **7**, 355–367.

[3] MCKINNEY, F., Studies in the retention of interrupted learning activities, *J. comp. Psychol.*, 1935, **19**, 265–296.

themselves equally well to open and closed situations, and they possess the additional merit of being relatively insusceptible to rehearsal. Crucial results would require the open forms to be presented, one, under conditions that preclude any possibility of tension and, two, under conditions in which a tension is operationally present. Actually, there seems no way of invariably insuring such a definitely unequivocal set of conditions, chiefly because the presence or absence of tension is so largely a function of the subject's attitude. We have tried to approach these respective conditions, one, by having subjects learn visual designs now open and now closed, and, two, by having them now interrupted and now uninterrupted while tracing visual designs.

## Method

The stimuli were irregular, novel designs, drawn in outline with black ink on 5″ × 3″ white cards. They ranged in area from 20 to 30 square centimeters. In Situation 1, 24 of these were presented initially to the subject; half of the 24 were open and half were closed. By 'open' is meant that approximately one-fifth of the periphery was left undrawn. To control variability in impressiveness of design, two alternate sets were constructed, so that the designs closed in one set were open in the other, and each of these sets was given to one of two comparable groups of subjects. Exposure was manual, for 2 seconds, with 2 seconds between stimuli. The subject was instructed to impress each design upon his mind so that he would be able to recognize it if he saw it again. After one presentation, the cards were mixed with 24 others, in an originally chance order that was constant for all subjects, and the whole 48 were presented for recognition. The interval between presentations was 5 minutes. One of three responses was required: 'yes,' 'no,' 'doubtful.' Thirty-two male college students acted as subjects.

In Situation 2, 20 closed figures, similar to those above, were presented one at a time to each subject, and he was required to trace the pattern through a sheet of thin rice paper. For half of the designs, he was interrupted by having his hand obstructed just before he reached the end; the other half of the designs he was allowed to complete. Again, alternate groups of subjects were interrupted on alternate designs in order to equate for varying impressiveness of figures. It was impossible to control time with any precision in this situation, and subjects were merely forbidden to loiter or to use haste. Nothing was said about learning the designs. The order of the cards, originally chance, was constant for all subjects, and on completion of the trial, they were mixed in constant arrangement with 20 others for recognition. Time between trials was five minutes. Responses were, as above, 'yes,' 'no,' 'doubtful.' The subjects were 32 male college students, and 22 children aged 10 to 12.

## Results and Discussion

In Table I are presented the data for Situation I. Reactions to the new stimuli are not given, since they are not

relevant to our problem. 'Doubtful' reactions are also omitted; their number can be determined by subtraction. Since there were 12 closed and 12 open stimuli in each series, and the subject had 3 possible reactions, 4 correct recognitions in each case would be expected by chance. Results on the two sets of stimuli, in which the conditions of openness and closure were alternated, have been pooled. This was necessary, for it was found, as expected, that some figures were much more easily recognized than others, the differences running all the way from below chance to 100 percent correct responses. (Incidentally, figures that were named by S almost invariably gave high recognitive scores.)

TABLE I

RECOGNITION OF OPEN AND CLOSED VISUAL FORMS

| | Recognized | | Unrecognized | |
|---|---|---|---|---|
| | C | O | C | O |
| Adults av. | 7.031 | 7.063 | 3.594 | 3.873 |
| (N 32) $\sigma$ | 2.685 | 2.738 | 2.320 | 2.254 |
| C.R. | .05 | | .84 | |

The purpose in Situation I was to compare retention of open and closed figures under conditions otherwise uniform. It is reasonable to suppose that, other factors being equal, an open figure would be less integrated than a closed figure; therefore, in Koffka's theory, it should have a lower survival value than the latter. But our data reveal no appreciable difference in this respect; the critical ratio is almost zero, and is, incidentally, in the opposite direction from that required by the theory. Possibly the factor of tension was not alike for the two figural conditions. It is known that open figures sometimes manifest a tendency toward closure. Assuming this tendency to be present, it may have induced in certain or all of our open stimuli a tension which compensated for their presumably greater instability. Or perhaps these stimuli sometimes gave rise to an attitude of annoyance which facilitated their retention. Unfortunately we obtained no introspective data relevant to this question. If increased tension

of some sort is the answer, it is a little surprising that it should balance so nicely the supposed loss from poorer organization. Perhaps a better answer is that our open figures were not deprived of enough of their form (about one-fifth) to affect appreciably their organization, and the subject therefore completed them perceptually quite easily and naturally.

The data for Situation 2 are given in Table II. 'Doubtful' reactions and reactions to the new stimuli are again not included; and responses on the 2 series, in which alternate sets of stimuli were interrupted, are pooled, as in Situation I. Since there were 10 completed and 10 interrupted tracings in each series, a chance score would be 3.333 correct responses.

TABLE II

RECOGNITION OF COMPLETELY AND INCOMPLETELY TRACED VISUAL FORMS

| | | Recognized | | Unrecognized | |
|---|---|---|---|---|---|
| | | C | I | C | I |
| Children (N 22) | Av. | 5.500 | 5.545 | 4.182 | 4.227 |
| | $\sigma$ | 2.961 | 2.296 | 2.609 | 1.965 |
| | C.R. | .057 | | .07 | |
| Adults (N 32) | Av. | 5.606 | 5.697 | 3.606 | 3.697 |
| | $\sigma$ | 3.625 | 3.873 | 2.954 | 2.034 |
| | C.R. | .099 | | .14 | |

Interruption is expected, normally, to cause a tension towards completion in the subject, and this tension is supposed to enhance retention by strengthening organization of the memory trace. But the data in Table II reveal no appreciable retentive superiority of the interrupted figures. The differences for both children and adults are in the predicted direction, but the critical ratios are so small as to make these differences meaningless.

Thus our results fail to support Koffka's theory. On the other hand, they should not be interpreted as constituting clear disproof of the theory. Our experimental conditions, we feel, left much to be desired, particularly in the matter of control of subjects' attitudes. Before positive conclusions can be reached, the effects of at least three different sets of conditions should be determined. (1) Effect of a greater in-

terval between presentation of the original stimuli and presentation for recognition. It may be that our interval was too short to allow differences in stability and organization to manifest themselves. A greater interval would permit interferences to weaken; although it would also allow more time for such complicating factors as assimilation and autonomous structural change of the forms. (2) Effect of a higher degree of learning. Perhaps our figures were not sufficiently impressed to begin with. Presumably one would have to stop short of overlearning in order to get an adequate measure of decay, but between this and our condition of one trial there are various degrees of learning that could be tried. (3) Effect of a more interesting task, that is, one the subject may be expected to want to complete. Tracing apparently had little or no appeal to most of our subjects, and interruption was therefore apt to be taken quite passively as a matter of course.

## Summary

Thirty-two adult subjects learned open and closed visual designs. Percentages of correct recognitions for the two conditions showed no differences. Thirty-two adults and twenty-two children were given the task of tracing designs. Half of the time they were interrupted. Percentage of correct recognitions was approximately the same for the interrupted as for the completed designs. The results do not support Koffka's theory of the effects of incompletion upon the survival of memory traces.

(Manuscript received January 11, 1940)

# EXTINCTION OF CONDITIONED PSYCHO-GALVANIC RESPONSES FOLLOWING TWO CONDITIONS OF REINFORCEMENT[1]

BY LLOYD G. HUMPHREYS

*Department of Psychology, Northwestern University*

An explanation of certain conditioning phenomena in terms of increments and decrements of excitatory potential for every reinforcement and non-reinforcement received has been questioned by several recent experiments. The present writer, for example, has reported results for the effects of random alternation of reinforcement in the acquisition and extinction of conditioned eyelid responses which contradict the hypothesis (**5**). Acquisition performance of a group of subjects given equal numbers of reinforced and non-reinforced trials was not significantly different from that of a control group given an equal number of trials all of which were reinforced. In extinction, however, the experimental group (50 percent reinforcement) gave responses of greater frequency and magnitude than the control group (100 percent reinforcement). The group with the larger number of reinforcements was thus less resistant to extinction, a result directly opposed to the prediction based upon the hypothesis in question.

The preceding findings have elsewhere been interpreted (**5**) in terms of preparatory set or expectancy. An adequate explanation is to some extent dependent, however, on knowledge concerning the generality of the findings. Thus it is the purpose of the present paper to report data for conditioned psycho-galvanic responses following 50 percent and 100 percent reinforcement. Whereas eyelid responses may be voluntary and reportable, psycho-galvanic responses are non-voluntary and non-reportable; yet the latter responses have been shown to be correlated closely with expectancy of stim-

[1] This research was conducted in the Department of Psychology, Institute of Human Relations, Yale University. The writer is indebted to the National Research Council for fellowship funds and to Professor Clark L. Hull who directed the research. Acknowledgment is made to Professor W. R. Miles for the use of a sound-resistant room and to Professor D. G. Marquis for the galvanometer used in the investigation.

ulating situations (3), (9). Comparability of results from eyelid and psycho-galvanic conditioning would substantiate the expectancy explanation. It would suggest, also, the non-voluntary character of the original eyelid results.

## Apparatus and Procedure

Data were available from a previous study (7) in which generalization following 50 percent and 100 percent reinforcement was investigated. The apparatus and procedure used were there reported more fully. Conditioned psycho-galvani c response were established to a tone of 1967 d.v., 40 db. above threshhold, which was followed at an interval of 400 ms. by an 'unpleasant' shock to the wrist. Other frequencies of equal subjective intensity, used to test for generalization, were 1718, 1311, and 1000 d.v. An experimental period consisted of the following: (1) two presentations of each of the four tones alone in a rotated order, or eight trials in all; (2) two presentations of the shock alone; (3) sixteen acquisition trials for 1967 d.v., all reinforced for one group of subjects (100 percent reinforcement), one-half reinforced for a second group (50 percent reinforcement); (4) repetition of step (1), or eight trials of the four tones alone. Since the order of presentation of the four tones in step (4) was systematically rotated, frequency of tone can be disregarded and the eight trials treated as an extinction series.

Trials occurred, on the average, at a 60-second rate. Subjects were undergraduates in Yale University and were paid by the hour. None had previously served in a conditioning experiment.

## Results

It has frequently been shown that the strength of conditioned responses is correlated with the strengths of the reflexes to the unconditioned and conditioned stimuli (1), (2), (10). This was found to be true of the present results. It was thus possible to match twenty subjects from the 50 percent and 100 percent reinforcement groups on the basis of initial sensitivity to the tones and to the shock (matching correlations of .94 and .96, respectively). This matching resulted in a correlation of .55 between the respective magnitudes of conditioned responses.

Figure 1 presents the magnitude of psycho-galvanic responses, in millivolts, for the trials preceding and following the differing conditions of reinforcement given the two groups of subjects. The lower, dotted lines represent the adaptation or extinction of the responses to the tones prior to reinforcement. Their striking similarity is of interest since it furnishes additional evidence for the initial comparability of the experimental groups. Had conditions of reinforcement remained the same, one would have expected no differences to appear following conditioning.

The conditioned responses following the two methods of reinforcement are represented by the upper, solid lines. It is seen that extinction now proceeds quite differently, the results being similar to those previously found for conditioned eyelid

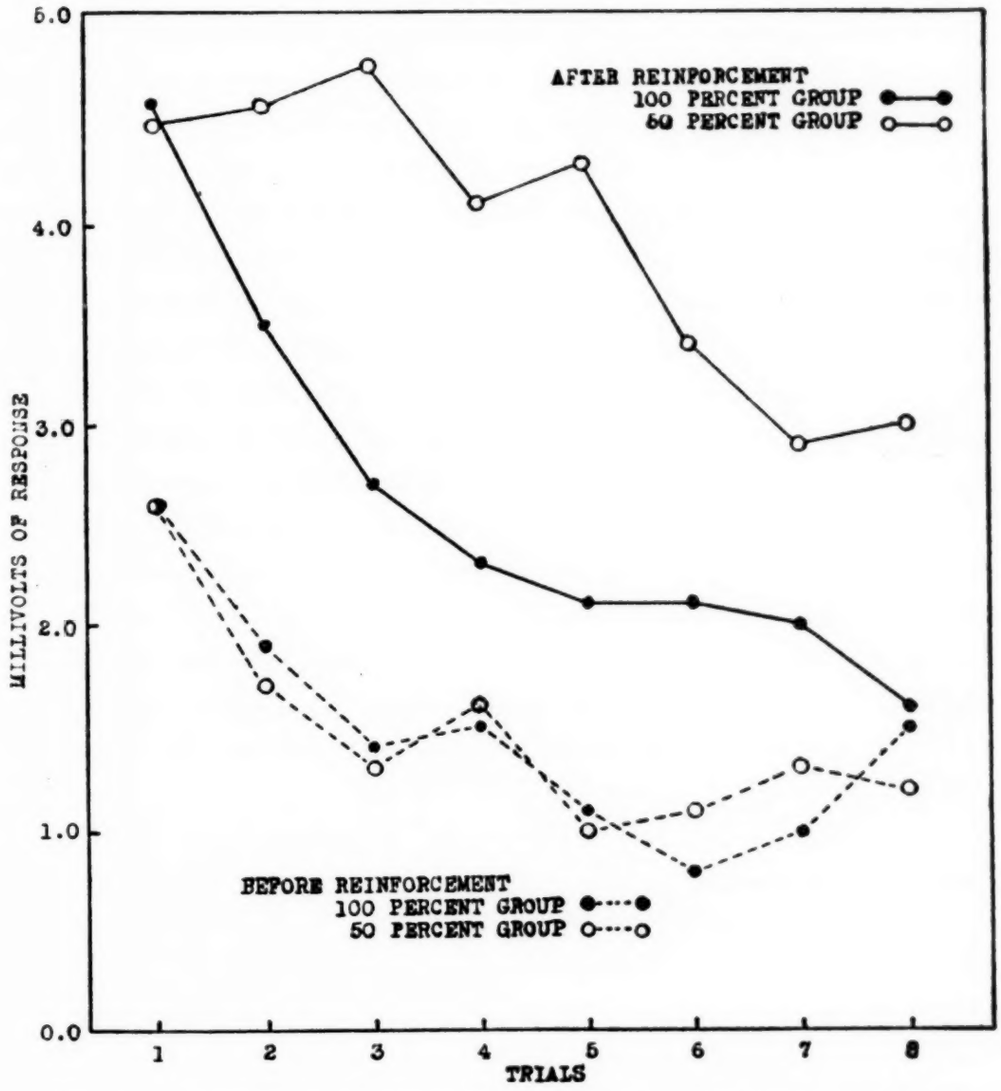


FIG. 1. Extinction of psycho-galvanic responses. Magnitude of responses is measured in millivolts. The lower, dotted lines represent trials prior to reinforcement. The upper, solid lines represent trials following reinforcement. For 100 percent reinforcement shock followed tone on every trial. For 50 percent reinforcement shock followed tone on one half of the trials. It is seen that the two groups of subjects, though initially comparable, perform differently following reinforcement. It is apparently more difficult for the 50 percent reinforcement group to form an expectancy of continued non-reinforcement.

reactions (**5**). Although one group had only eight reinforcements in sixteen trials as compared to sixteen reinforcements for the second group, the magnitude of conditioned responses

is comparable at the end of acquisition (first extinction trial). As extinction proceeds, however, the 100 percent group manifests the usual negative acceleration, while the 50 percent group rises to the third trial and drops off more slowly thereafter. The *t* test for the significance of this extinction difference shows that the probability is less than .01 that a difference as large could have arisen by chance (*t* equals 3.23).

The preceding results cannot be explained by the operation of increments and decrements due to reinforcement and non-reinforcement.[2] They are congruent, however, with an explanation in terms of expectancy, *i.e.*, that acquisition and extinction of the conditioned responses are dependent on the formation of expectations or 'hypotheses' concerning reinforcement and non-reinforcement. Comparable acquisition performance in spite of 50 percent reinforcement may be explained on the basis that the noxious character of the unconditioned stimulus compensates for a considerable degree of uncertainty regarding its appearance. The extinction results are apparently due to difficulty in forming an hypothesis of continuous non-reinforcement after discontinuous reinforcement in acquisition. This assumption has been verified for verbal expectations (**6**) and is also supported by the results of Mowrer (**9**) and of Cook and Harris (**3**).

Extinction following 50 percent reinforcement is at least phenotypically similar to the extinction following continued, massed reinforcement reported by Hovland (**4**). The present results cannot be explained by the 'inhibition of reinforcement' concept, however, since the initial rise in the extinction curve is associated with the smaller number of reinforcements. An expectancy explanation for distributed practice phenomena, on the other hand, seems also to be ruled out, as Humphreys, Miller, and Ellson (**8**) found no relationship between inter-trial time intervals and the acquisition and extinction of verbal expectations.

[2] It should be noted that the factor of distributed reinforcement in the 50 percent group was controlled in the previous eyelid conditioning (5). Since extinction proceeded in much the same fashion whether reinforcement had been massed or distributed, this control was omitted in the present investigation.

## Summary and Conclusions

Conditioned psycho-galvanic responses were established in forty human subjects to a tone of 1967 d.v. under two conditions of shock-reinforcement. One group of twenty subjects was given sixteen paired presentations of the tone and shock (100 percent reinforcement); the remaining twenty subjects were given sixteen presentations of the tone, eight of which were paired with shock in a random order (50 percent reinforcement). From the results it can be concluded:

1. Performance at the end of acquisition is approximately equal under the two conditions of reinforcement. This finding parallels the results from previous eyelid conditioning.

2. Extinction following 100 percent reinforcement exhibits the usual negative acceleration, that following 50 percent reinforcement an initial rise and less rapid decline. This finding is similar both to the eylid conditioning results and to those for verbal expectations.

3. The results here reported are congruent with the explanation in terms of preparatory set or expectancy advanced in a preceding paper.

(Manuscript received January 3, 1940)

### References

1. Campbell, A. A., The interrelations of two measures of conditioning in man, *J. exper. Psychol.*, 1938, **22**, 225–243.
2. Campbell, A. A., and Hilgard, E. R., Individual differences in ease of conditioning, *J. exper. Psychol.*, 1936, **19**, 561–571.
3. Cook, S. W., and Harris, R. E., The verbal conditioning of the galvanic skin reflex, *J. exper. Psychol.*, 1937, **21**, 202–210.
4. Hovland, C. I., 'Inhibition of reinforcement' and phenomena of experimental extinction, *Proc. Nat. Acad. Sci.*, 1936, **22**, 430–433.
5. Humphreys, L. G., The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions, *J. exper. Psychol.*, 1939, **25**, 141–158.
6. Humphreys, L. G., Acquisition and extinction of verbal expectations in a situation analogous to conditioning, *J. exper. Psychol.*, 1939, **25**, 294–301.
7. Humphreys, L. G., Generalization as a function of method of reinforcement, *J. exper. Psychol.*, 1939, **25**, 361–372.
8. Humphreys, L. G., Miller, J., and Ellson, D. G., The effect of the intertrial interval on the acquisition, extinction, and recovery of verbal expectations. (To appear.)
9. Mowrer, O. H., Preparatory set (expectancy)—a determinant in motivation and learning, *Psychol. Rev.*, 1938, **45**, 62–91.
10. Pavlov, I. P., *Conditioned Reflexes*. London: Oxford University Press, 1927.

# STUDIES IN THERMAL SENSITIVITY: 14. PART-WHOLE RELATIONS IN SERIATIM WARM-MAPPING

BY WILLIAM LEROY JENKINS,

*Lehigh University*

In a previous study,[1] it has been shown that there are highly consistent relations between part and whole seriatim scores when the same skin area is mapped alternately, at 17° C., with two different stimulator sizes. In the present study, the analogous problem has been investigated for warmth.

## Procedure

The method was essentially the same as that employed in the study with cold. Subjects worked in groups of three, each acting in turn as experimenter, subject and recorder. A checkerboard pattern of suitable squares or rectangles, stamped on the volar surface of the left forearm, was mapped alternately with the two sizes, with suitable rest periods between, until a total of 12 mappings had been completed. The report categories were: 0 for neutral, 1 for weak, 2 for medium and 3 for strong warmth. Inlet temperature was maintained at 44° C. in all experiments—a level which preliminary experiments indicated was unlikely to cause pain.

The chief difference from the corresponding cold-mapping experiments lay in the temporal spacing of the successive mappings. Since preliminary work [2] had indicated that warm-mapping was peculiarly open to artifacts arising from adaptation and vascular changes, the time between successive mappings was lengthened to 15 minutes, and the total experimental session was divided into two parts with a two-hour rest period between.

## Results

The data in general show such a remarkable similarity to those for cold that their presentation may be greatly abbreviated. Table 1 gives the distribution of reliability coefficients between the whole scores and the square-root function of part scores for the three size comparisons, and for the odd-even correlation of 6 mappings with a 1 mm diam. stimulator.

[1] Jenkins, W. L., Studies in thermal sensitivity: 12. Part-whole relations in seriatim cold-mapping, *J. exper. Psychol.*, 1939, **25**, 373–388.

[2] Jenkins, W. L., Studies in thermal sensitivity: 10. The reliability of seriatim warm-mapping with untrained subjects, *J. exper. Psychol.*, 1939, **24**, 439–449.

The general level of reliability is somewhat lower than for seriatim cold-mapping, slightly less than half the group reaching the required standard of .80. With the 1 mm diam. stimulator, discrimination broke down completely for the majority of the subjects and only 11 out of 92 reached .80 or higher. Consequently, no size comparisons were made involving this stimulator. In all of the following analyses, only the figures from those subjects showing reliability coefficients of .80 or greater have been used.

TABLE 1

| | Less than .59 | .60 .64 | .65 .69 | .70 .74 | .75 .79 | .80 .84 | .85 .89 | .90 .94 | .95 1.00 | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 × 1 2 × 2 | 7 | — | 3 | 3 | 2 | 3 | 6 | 3 | — | 27 |
| 2 × 2 2 × 4 | 10 | 3 | 2 | 1 | 4 | 8 | 6 | — | 2 | 36 |
| 3 × 3 3 × 6 | 4 | — | 1 | 3 | 8 | 3 | 4 | 8 | 1 | 32 |
| 1 mm diam. | 53 | 3 | 5 | 6 | 14 | 7 | 4 | — | — | 92 |

Figure 1 shows in condensed form the essential part-whole relations. In part *A*, the circles show the median whole scores for cases where the two part scores are identical. (For clarity, the part scores have been grouped into 7 category levels.) The heavy line is *not* fitted to these data, but is the basic curve earlier developed for *cold*.[3] It is evident that the law for summation from identical scores is the same for cold and for warm, except for a deviation at the '3' level, which will bear further investigation. Part *B* shows plainly that the maximum of part scores does not fit the basic curve, whole scores being consistently less than the maximum of part scores. Parts *C* and *D* demonstrate that there is a very good fit of the square-root function of part scores to the basic curve, for various part score differences and for the three size comparisons used.

[3] JENKINS, W. L., Studies in thermal sensitivity: 12. Part-whole relations in seriatim cold-mapping, *J. exper. Psych.*, 1939, **25**, 377.

Figure 2 shows seriatim maps made with the 1 mm diam. stimulator for six subjects whose reliability coefficients ranged from .84 to .88. As with cold, there are hills and valleys of sensitivity, but isolated positive squares ('warm spots') surrounded by zero squares are conspicuously absent.

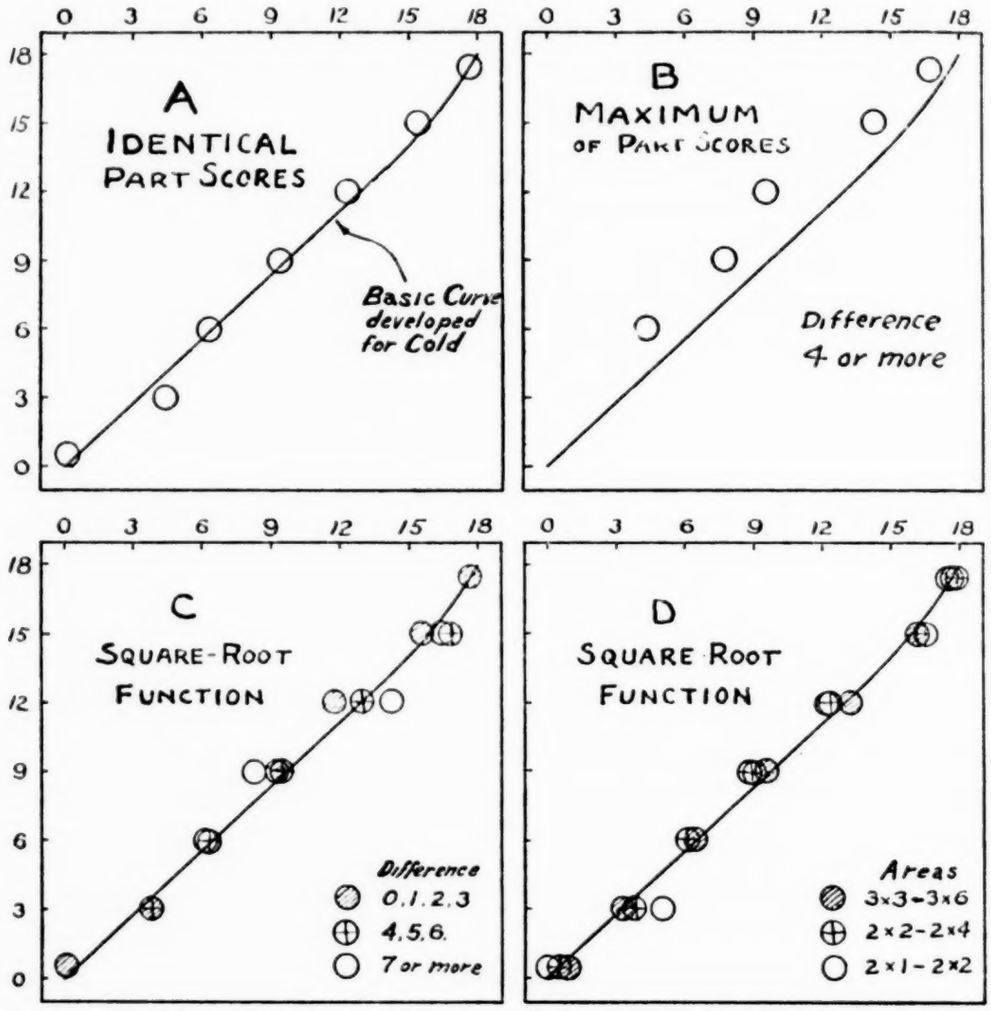


FIG. 1. Ordinates: functions of part scores. Abscissae: whole scores. Circles: medians of whole scores.

## DISCUSSION

The data for part-whole relations are so similar for warm and cold, that it seems unnecessary to repeat here the detailed discussion of the earlier paper.[4] The facts of warm sensitivity, like those of cold, cannot be adequately explained in

[4] *Op. cit.*, 381–388.

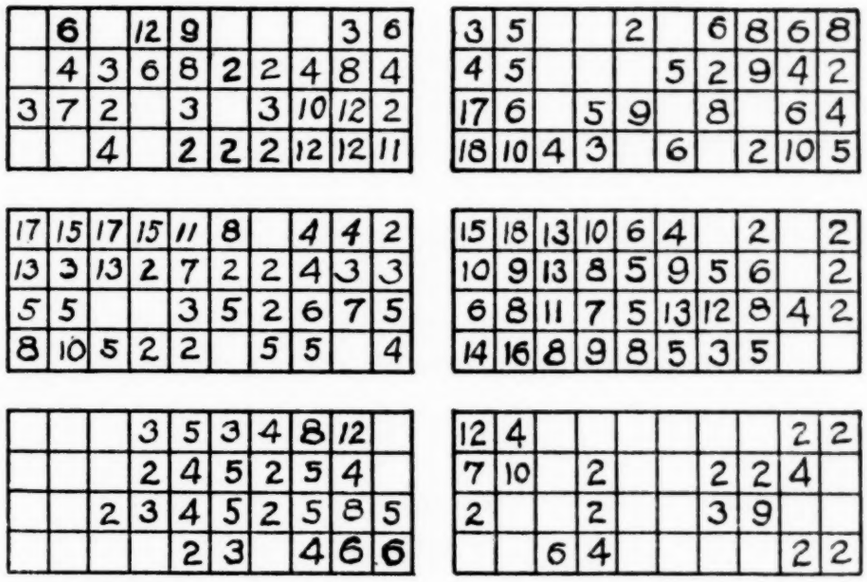


FIG. 2. Seriatim maps with 1 mm diam. stimulator at 44° C. Each square represents 1 mm². Reliability coefficients from .84 to .88.

terms of the traditional theory of 'warm spots,' nor by assuming varying thermal conduction to account for the differing score levels. The *concentration theory*, however, can be equally well applied to warm sensitivity.

One objection, advanced since the publication of the earlier paper, can be disposed of briefly. In part *B* of Fig. 1, it is shown that whole scores are consistently less than the maximum of differing part scores. It has been suggested that this reduction might be caused by the greater *tactile interference* from the double-sized stimulator. If this were true, however, the same sort of tactile interference should be present when the part scores are identical. Part *A* of Fig. 1 shows that with identical part scores, the median whole score is *greater* than the part scores. Hence, tactile interference cannot account for the relationship.

More serious is the contradiction between these results and those obtained by Hardy and Oppel with radiant stimulation.[5] These investigators showed that the *threshold* for radiant

[5] HARDY, J. D. AND OPPEL, T., Studies in temperature sensation. III. The sensitivity of the body to heat and the spatial summation of end-organ responses, *J. Clin. Invest.*, 1937, **16**, 535–540; IV. The stimiulation of cold sensation by radiation, *J. Clin. Invest.*, 1938, **17**, 771–778.

warmth and cold varies as an inverse function of stimulation size up to approximately 200 $cm^2$—in short, very marked summation effects. In contrast, our seriatim mapping studies indicate summation to be a very minor influence in determining seriatim scores.

The contradiction, however, may be more apparent than real. Hardy and Oppel worked entirely with radiant stimulation; the seriatim mapping studies have been made entirely with contact stimulation. Hardy and Oppel measured only threshold intensities; our studies have been always with definitely supra-liminal intensities. Certain technical difficulties lie in the way of seriatim mapping with radiant stimulation, or of threshold determinations with contact. It is hoped, however, that further research may serve to clarify the situation, since both sets of results appear to be soundly established.

## Summary

The part-whole relations found in seriatim warm-mapping are virtually identical with those earlier reported for cold. Aside from a minor effect of size as such, whole scores are expressed as a square-root function of part scores. Seriatim maps with a 1 mm diam. stimulator show no isolated 'warm spots.' Neither the traditional 'warm spot' theory, nor the thermal-conduction modification of it provides an adequate explanation of the experimental facts. Warm sensitivity, like cold sensitivity, is best explained in terms of a *concentration theory*, *i.e.*, that sensitivity is determined by the concentration of minute receptors.

(Manuscript received December 19, 1939)

# AN APPARATUS AND METHOD FOR STIMULATING, RECORDING AND MEASURING FACIAL EXPRESSION

BY JOHN G. LYNN, M.D.*

*Neurological Institute, New York*

Since 1935 the writer has conducted a series of studies on contrasting aspects of the dynamic patterns of involuntary facial expression and their relation to general patterns of behavior. The subjects of these investigations have consisted of approximately 1,200 normal and abnormal individuals. Reports of this work have been and are being published.[1]

In the course of these studies, it has been found possible to discard inexact, subjective analysis of involuntary facial reactions, and to substitute in its stead a technique for exact and objective measurement.

An apparatus, the facial cinérecorder, has been devised, and has been used successfully on 108 patients at the Psychiatric Institute, N. Y. C. This set-up combines a continuous automatic sound movie projector with a concealed moving picture camera so that facial expression is first auditorially and visually stimulated and then photographed under standard conditions without the subject's knowledge. The developed cinematographic record can later be subjected to an analysis in which the dynamic patterns of expression are translated first into graphic and finally into quantitative numerical representations. The type of moving picture shown to the subject will depend on the particular emotional reac-

* The author was aided in the development of a suitable sound and telephone attachment for the continuous automatic projector by Mr. Harry B. Jones, research engineer for Advertising Projectors, Inc., 10 West 33rd Street, New York, N. Y.

[1] LYNN, JOHN G. AND DORIS R., Face-hand laterality in relation to personality, *J. abnorm. and soc. Psychol.*, July, 1938, **33**; and Laterality of involuntary mimetic smiling: an indicator of aggressive and retiring tendencies. A cinematographic analysis; Presented in brief form at the 29th annual meeting of the American Psychopathological Association, Atlantic City, June 6th, 1939. To be published.

tions desired. As involuntary mimetic smiling and laughter were being studied in the forementioned investigation, Walt Disney films were used.

Figures 1 and 2 show, from slightly different angles, the construction and use of the facial cinérecorder. The apparatus has been built in four separate sections, each with a carrying handle, so that they can be transported easily. These sections are respectively, I, the facial aperture box (1)

FIG. 1. The facial cinérecorder in use with side covering dropped to show internal construction.

II, camera-screen-light unit (17, 11, 18), III, continuous automatic projector (21), and IV, photo-electric sound unit (15). When assembled, sections I, II, and III lock together to form a rigid unit with over-all dimensions of $57 \times 18\frac{1}{4} \times 17\frac{1}{2}$ inches. Section IV, the sound unit, rests on top of sections II and III. The facial aperture box (1) is attached to the camera and screen unit by metal tie-rods (5) so that the subject's face is always three feet from the camera lens (17).

The space between the face box and the camera box is covered with black photographer's cloth in order to minimize outside stimulation. This covering has one side unhooked and dropped down in the photograph (Fig. 1) and is omitted in the diagram (Fig. 2). The whole apparatus is placed on a long table.

The subject sits on an adjustable stool in front of the facial aperture box. He is told that the purpose of the experiment is to test his memory of the Mickey Mouse movie to be seen and heard. This memory test serves as a necessary cover-up for the real purpose of the experiment. It gives the

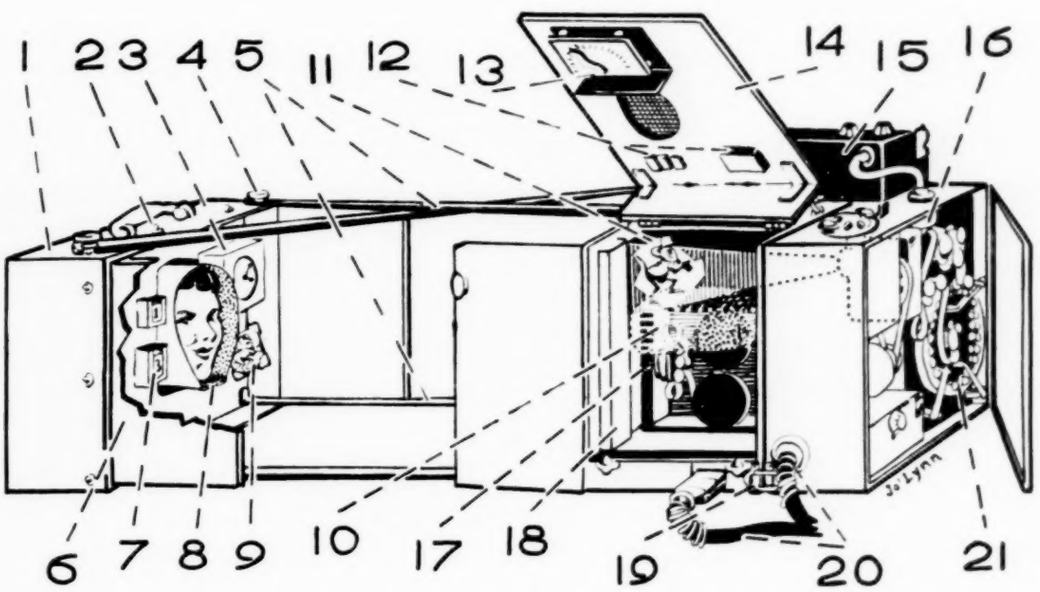


FIG. 2. Diagram of the facial cinérecorder.

subject a plausible reason for putting on the earphones and placing his face in the adaptable opening (8) in the exact position to be photographed. While he is watching the movie on the 5 × 7 inch daylight translucent screen (11) and laughing at Mickey's antics, a concealed high-speed camera (17) takes his picture through one opening in a duplex electric outlet situated just underneath the screen. One of these two outlets is functional and serves as the connection for the electric time-clock (3). The other outlet is merely a camouflage for the camera lens. On either side of the screen and beneath it, rows of diffuse General Electric lumiline tubes (18)

throw sufficient light on the subject's face for it to be photographed easily on ultra-super-sensitive film with the lens at F.1.5 and with a speed of 40 frames a second. These lights are not sufficiently bright or glaring to annoy the subject or to make him suspicious. In order to prevent the heat from the enclosed lights from annoying the subject or damaging the film record, two electric blowers are used to suck air around the lumiline tubes into the camera box and to blow it out of apertures on the two sides. The position of one of these blowers is shown by a ghost outline (10) on Figure 2.

Besides photographing the subject's facial reactions, the camera also records the face of a time-clock (3) which measures 1/100 second intervals, a mirror image of the picture on the screen inducing the reaction, and various index cards (7) displaying the subject's number, the date, etc.

The moving picture on the screen is projected from the rear from the continuous automatic projector (21). The sound apparatus (15, 16) which has been built for this projector, serves the triple purpose of intensifying the humor of the picture, of providing earphones which keep the subject's head in place, and of hiding the intermittent noise of the camera.

The operator not only controls the projector, camera, lights, and sound effects, but he can also talk to the subject through a hand-telephone, listen to the sound effects through his own earphones, and can watch the subject through an opening in the top of the black cloth covering or through a periscope. In order to facilitate smiling and to conserve film, it was found advisable in the experimentation at the Psychiatric Institute, to keep the top of the black cloth covering open so that the operator could smile at as well as speak to the subject through the phone in a standardized manner at designated points in the movie. (See Fig. 1.) Immediately preceding the injection of this additional stimulation, the camera was turned on, ready to catch any changes in the patient's expression. Under these circumstances, smiles were obtained from 93 percent of the 108 subjects already studied with the facial cinérecorder. As far as could be determined, only one patient suspected that his picture was being taken.

## Method of Mapping Facial Expression from Moving Picture Records

At some later date, the developed motion picture record of each patient's expression (See Fig. 3) is projected onto the back of a $19\frac{1}{2} \times 26\frac{1}{2}$ inch translucent screen forming part of the surface of an architect's drawing board. A movie projector on the floor is so arranged in conjunction with the drawing-board screen that every distance on the screen is exactly double the actual size. This magnification helps to

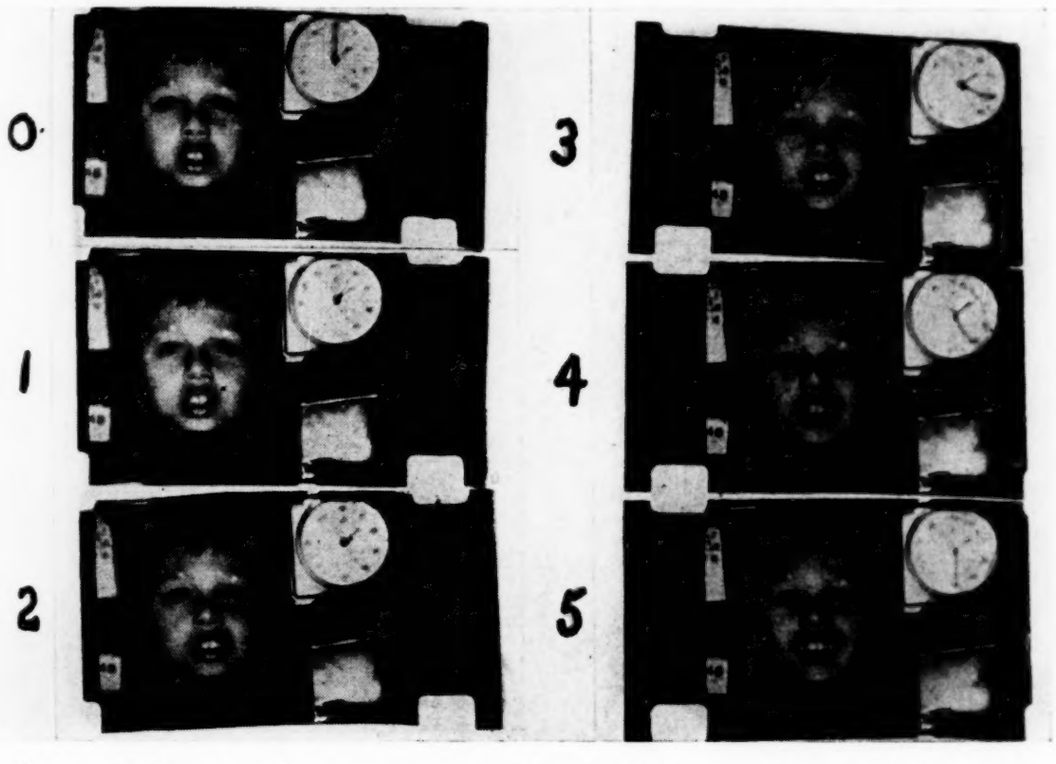


Fig. 3. Subject's smile as photographed by facial cinérecorder. Frames at 1/10 second intervals.

reduce errors incidental to operations of measurement. A transparent tracing paper is fixed against the drawing-board screen and the movie projector is manually operated by an assistant. At the beginning of a smile, the projector is stopped and relatively invariable points of reference on the face, such as eye-corners, eyeglasses, upper teeth, hair line and sides of the face over the zygomatic arches are traced off. These serve collectively as a stable frame of reference by which the tracing paper can be kept in a standard position on

the face regardless of the amount of activity of the muscles of expression. Next, the passive position of the outer corners of the mouth at the muco-cutaneous junction are indicated by two dots (See o and o in Fig. 4). Commencing at the start of a smile and using the stable points of reference and the photographed time-clock as guides, the projector is slowly operated and stopped regularly at 1/10 second intervals in order to record any change in the position of the mouth corners. (See Figs. 3 and 4.)

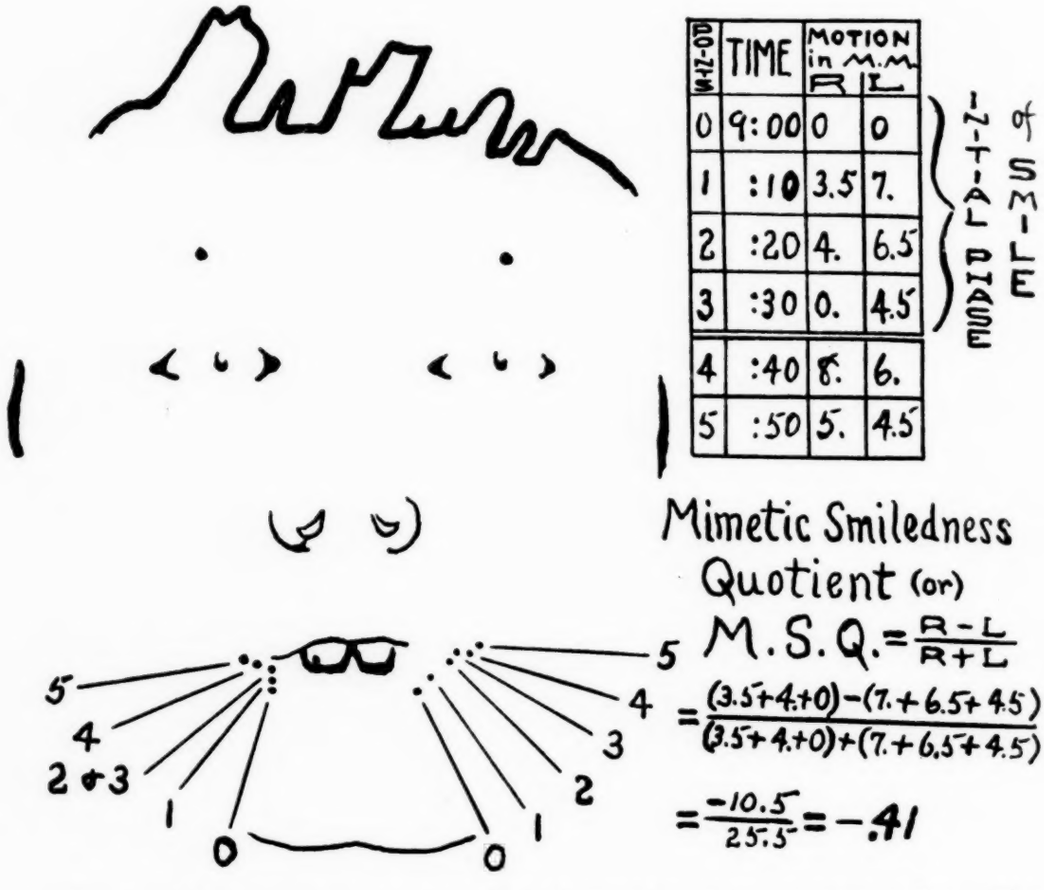


FIG. 4. Diagram illustrating method of mapping and measuring smiles. The index of initial laterality (Mimetic Smiledness Quotient) used in the Psychiatric Institute study is represented.

The reliability of this method of mapping was determined by engaging the services of a disinterested person. She independently remeasured twenty of the smiles chosen at random. The measurements from her maps of the amount of activity of the right and left sides of the face respectively were correlated with the amount of activity as shown by our

maps of the same smiles on the same individuals. A +.98 coefficient of reliability was obtained for this method of mapping and measuring the dynamic patterns of involuntary mimetic smiling or laughter.

## Measurement of Facial Activity from Maps of Facial Expression

From the maps of the paths of motion of the mouth corners in conjunction with simultaneous time-clock readings, one can obtain by direct measurement reliable numerical values of activity and relative rates of motion. The forms assumed by these indices will depend on what phase of expression is being measured. In the studies so far made, lateral hyperactivity of the mouth corners during involuntary smiling was being measured. The initial phase of the smile was of first importance and therefore the index represented the differential rates of activity of the mouth corners during the early phase of the upward motion of the smile. The reliability of the particular index that was used in our study was found to be +.61 with a P.E. of ±.05 between smiles occurring within a three-minute interval. Twenty-one of the subjects were retested three weeks later and the reliability coefficient between the average of the indices of smiles recorded at the first and at the control test proved to be +.59.

Indices could also be established for the differential rates of motion or for the differences in absolute distances travelled by the two sides during the smile as a whole. The amount of right or left temporal lead could be determined with even greater ease and accuracy. For other facial reactions than the smile, other mathematical represenatations may be used. As an example, eyebrow and forehead movements can be efficiently recorded for measurement by the use of the facial cinérecorder.

In order to examine eyebrow movements in the Psychiatric Institute study, the patient was first given a pseudo eye-test in which spots of white adhesive were placed on his eyebrows directly over his pupils when looking straight ahead. He was asked to find the mirror reflections of these spots

through a hole in a card. The standard Hole-in-a-Card tests for eye-dominance were finally administered without the use of the mirror. These latter tests served not only to give information on eye-dominance, but also to divert the patient's attention from the presence of the eyebrow spots; so that he invariably forgot them when introduced immediately to the facial cinérecorder. The small spots on the eyebrows formed indispensable reference points for the mapping and measuring of eyebrow and forehead activity.


## Resumé

An apparatus, the facial cinérecorder, has been devised which combines a continuous automatic sound movie projector with a concealed moving picture camera so that facial expression is first auditorially and visually stimulated and then photographed under standard conditions—all without the subject's knowledge.

Next the developed cinematographic record is projected onto a translucent screen and subjected to an analysis in which the dynamic patterns of facial expression are first traced off graphically and then measured and translated into numerical representations.

The facial cinérecorder, plus the forementioned technique of analysis of the film records, has made it possible to replace subjective judgments of facial expression by objective measurements. A +.98 coefficient of reliability was obtained when this quantitative method was used in mapping and measuring the dynamic patterns of mouth corner activity during involuntary smiling.

By changing the nature of the stimulus situation, many varieties of emotional reaction other than smiling can be induced and studied.

Other obvious uses of the new apparatus and method would be in the investigation of the kinds and degrees of loss and recovery of function occurring in both involuntary and voluntary facial palsy. Furthermore, eyelid activity, eye movements, forehead and eyebrow motion can just as easily be made the subject of measurement as mouth corner activity.


(Manuscript received January 15, 1940)

# AN IMPROVED STABILIMETER METHOD FOR RECORDING ACTIVITY IN LABORATORY ANIMALS

BY KARL U. SMITH

*The University of Rochester*

In making observations on the spontaneous activity of animals after experimental lesions of the cerebral cortex, the apparatus described below has been found to have certain advantages over other types of stabilimeters in successfully recording the activity of typical laboratory subjects such as the rat, guinea pig, cat, and monkey. Some of the features of a stabilimeter-type activity cage which are essential for long-time recording include durability, ease of adjustment of the apparatus, fair sensitivity, and easy dismantling for cleaning, readjustment, etc. It is believed that the apparatus described possesses these features, and, in addition, can be calibrated to an approximate degree so that the magnitude of the movements of the animal may be estimated.

Figure 1 illustrates the arrangement of the stabilimeter

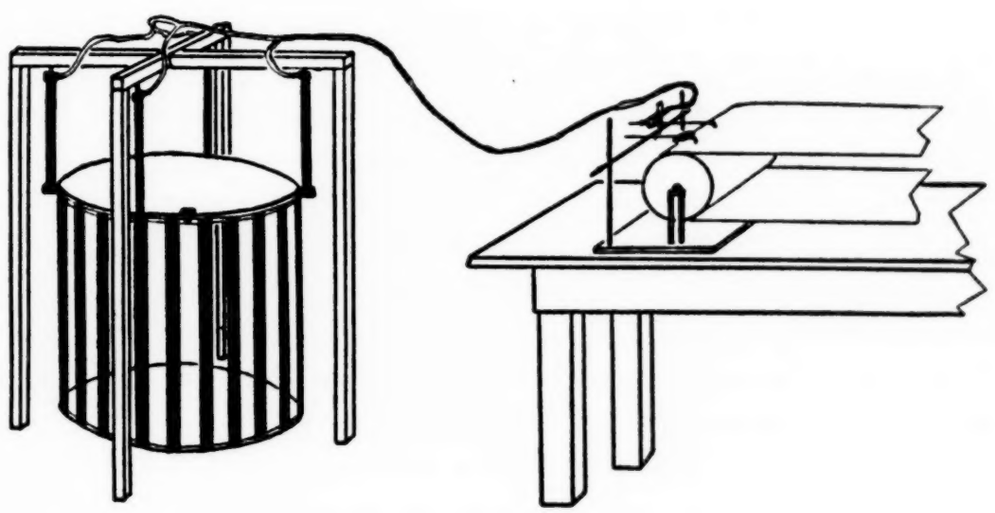

FIG. 1. The arrangement of the stabilimeter and recording system for observations on the behavior of cats. The pneumograph supports for the activity cage, shown to the left, are drawn in black.

and a recording system as used in determining the nature of spontaneous activity in cats before and after lesions of the frontal areas of the cortex. The stabilimeter is shown to the left of the figure while the kymograph and ink-recording system are represented toward the right. The cage of the stabilimeter consists of a cylinder, 28 inches in diameter and 22 inches high, enclosed by narrow plywood strips, ¼ inch thick and ½ inch wide. The strips are mounted one inch apart on two plywood rings, which also support the plywood discs which form the top and bottom of the cage. The top disc is attached by means of hinges and thus can be swung upward to admit the animal into the cage.

The cage is suspended from supporting cross-beams by means of four pneumographs, somewhat less sensitive than the typical pneumographs used for registering respiratory movements. These pneumographs are constructed of ¾-inch wire springs spun from No. 18 spring-brass wire, enclosed in ¾-inch thin-wall band-stock rubber tubing. The ends of the springs are attached by means of small pins to the inside of brass-pipe caps, and the tubing is stretched and wired over the outside edge of the caps. The cap at the bottom of each pneumograph is attached to the cage by means of a hook, soldered to its tip, while the cap at the top of each pneumograph is drilled and fitted with a 3-inch piece of ¼-inch brass tubing. The pneumographs are held in place at the supporting beams by clamps.

Leading from the air outlet provided by the piece of brass tubing at the top of each pneumograph is a length of small-bore rubber tubing. The four pneumograph outlets are connected together by means of T-shaped glass rods and brought to a common outlet which leads to a metal bellows on the ink-recording kymograph. A metal capillary lever, which is driven by the movement of the bellows, records all movements of the cage.

Extra-flexible hydron metallic bellows are used in the set-up in place of the ordinary rubber tambour for recording changes in air pressure in the pneumograph system.[1] The

[1] Clifford Manufacturing Company, Boston, Mass.

bellows used here is a Type F, 6-ply bellows, $2\frac{7}{16}$ inches in diameter. The head of the bellows is attached to a support, which holds it in place, and is connected directly to the common outlet of the pneumograph system. A small metal guide, fixed to the middle of the closed end of the bellows, is fitted to the capillary lever and serves to move the lever when the bellows expands or contracts. Ink is fed into the capillary lever from an inkwell mounted on the support to the side of the bellows.

The apparatus described here is sufficiently sensitive to record any or all of the slight movements which the cat normally makes during the waking or sleeping state. Experimentation has shown that the four-point support by the pneumograph system provides a more sensitive recording system than either a three- or two-point support. At the maximum sensitivity of the system, movements which the animal makes in licking itself produce ¼- to 1-inch deflections on the kymograph record, while gross movements of the animal are recorded as 3- to 5-inch deflections. Sensitivity may be modified by changing the position of the driving point of the bellows along the length of the capillary lever, or by introducing slight leaks in the pneumograph system. This latter adjustment is used to eliminate the recording of the slower swinging movements of the cage, and to introduce a constant base line in the activity record.

Record A in Figure 2, which illustrates the recording of spontaneous activity, was secured on an animal in which the frontal lobes of the cortex had been removed. Record B in this figure demonstrates the use of the apparatus in recording the development of a conditioned response. In this situation the conditioned stimulus was a bell, the unconditioned stimulus, a shock to both hind paws, and the response, general bodily movements, recorded as shown. These records were obtained when the sensitivity of the recording system was adjusted to approximately half its maximal value and when a small vent was introduced into the air system.

A similar but smaller stabilimeter has been constructed for studies on spontaneous activity and conditioning in guinea

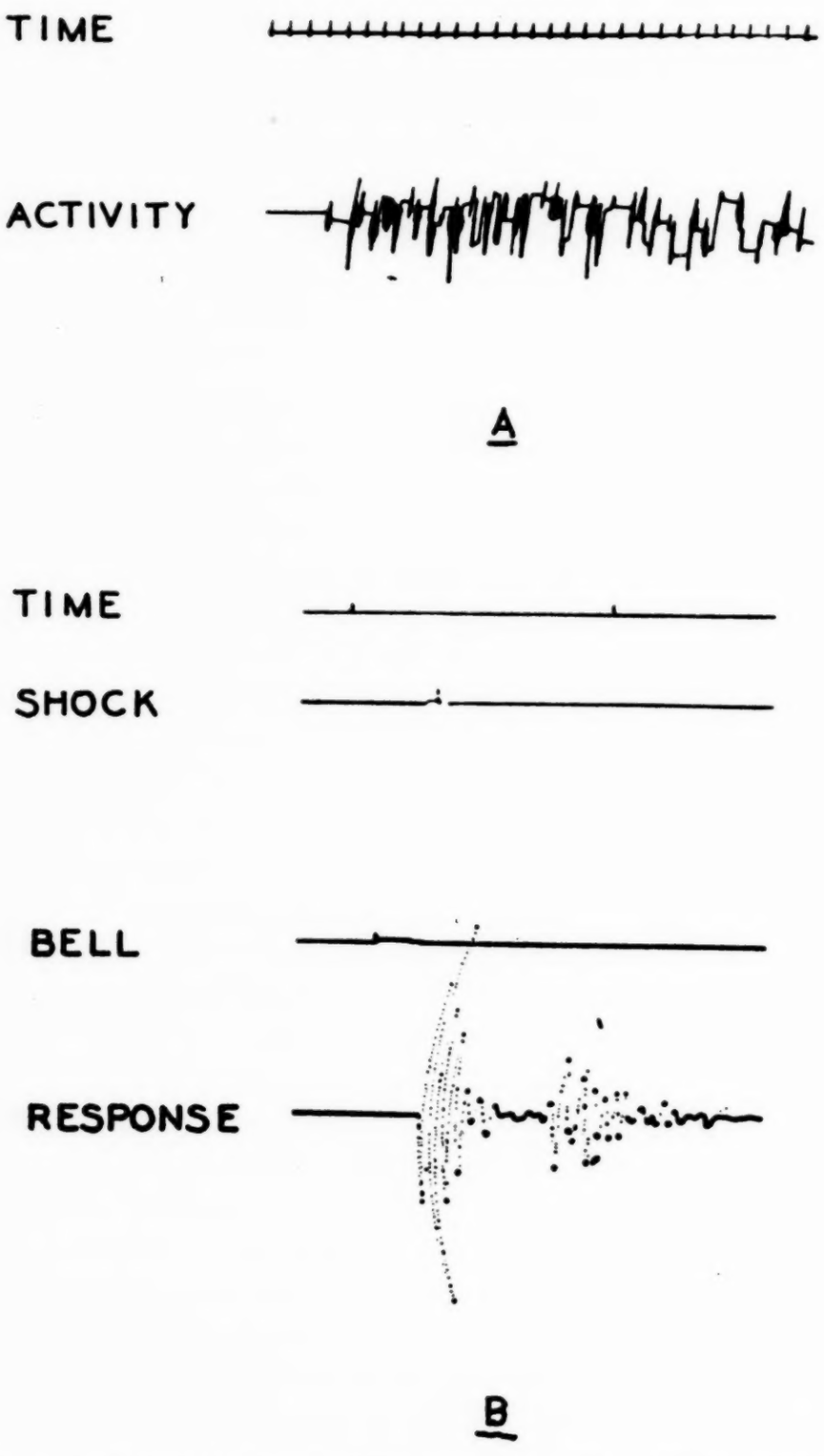


FIG. 2. *A*: Record of the spontaneous activity of a cat in which the frontal lobes of the cortex had been removed. The time line is in minutes. *B*: Record of the response of a normal cat to the unconditioned stimulus, a shock to the two hind paws, during the establishment of a conditioned response to a bell. The time line represents 15-second intervals.

pigs. This cage, which consists of a small wooden frame enclosed on its top and sides by sheet celluloid, is supported by pneumographs of greater diameter and sensitivity than those employed with the cat.

The type of stabilimeter described here has a number of characteristics which make it superior to the tambour-supported activity cage which is frequently employed. The present apparatus is very easily constructed and will withstand any degree of hyperactivity of the animal. Air-leaks within the pneumograph system are easy to control and repair since the use of rubber tambours has been dispensed with. The use of a metal bellows in the recording system permits a rough calibration of the extent of displacement of the recording lever in terms of the extent of the movements of the animal within the cage. This is impossible when perishable rubber tambours are used in the recording system. Finally, the pneumograph supports provide a much more sensitive recording system than is obtainable when tambour bases are used.

(Manuscript received December 21, 1939)